Adoption of the Internet by Commercial Establishments: Urban density, Global Village and Industry Composition

March 2003

Chris Forman, Avi Goldfarb, and Shane Greenstein*

Abstract

Two opposing views have been argued on the relationship between Internet technology and economic agglomeration. One view, which we term global village theory, asserts that Internet technology helps lower communication costs and break down geographic boundaries between firms. The other view, labeled urban density theory, argues that the Internet follows a traditional pattern of diffusion—diffusing first through urban areas with complementary technical and knowledge resources that lower the costs of investing in new frontier technology. In this paper, we offer hard evidence on factors influencing the dispersion of Internet technology to businesses. On the one hand, we find no evidence for urban density theory in the diffusion of basic access and participation in the Internet network. We do find some evidence supporting global village theory for diffusion along this dimension. On the other hand, we find that the pattern of adoption of frontier Internet technologies supports urban density theory. Our results reject the concept of an urban-rural digital divide in participation in the Internet network. Even for frontier technologies, the pre-existing distribution of industries determines most of the differences between locations. Consequently, policy aimed at increasing adoption in laggard areas may be misguided.

---

* Respectively at Carnegie Mellon University, University of Toronto and Northwestern University. We thank Tim Bresnahan, Steve Klepper, Roger Noll, Alicia Shems, Scott Stern, and Manuel Trajtenberg for comments. We also thank Harte Hanks Market Intelligence for supplying data. We received funding from the Kellogg School of Management and the GM Strategy Center and seed funding from the National Science Foundation and Bureau of Economic Analysis. All opinions and errors are ours alone.

1. Introduction

General Purpose Technologies (GPTs) involve high fixed costs in invention and low marginal costs in reproduction. Almost by definition, GPTs have a big impact if and when they diffuse widely, that is, if they raise the marginal productivity of a disparate set of activities. As a practical matter, "disparate" means a great number of applications and industries performed in a great number of locations.[1] As with most GPTs, the Internet was a malleable technology when it first commercialized in 1992, and it needed to be adapted for commercial use. Adaptation was necessarily a local economic activity, resulting from the combination of the demands of business establishments and the supply constraints of markets for Internet technology infrastructure and services.

More pertinent to our study, GPT theory predicts that the same technological opportunity (i.e., the commercialization of the Internet across the United States) does not result in the same commercial experience for all establishments in all locations. However, a lack of systematic data and formal theory has made it difficult to characterize variance in Internet experience among commercial users. With a few notable exceptions, there has been little systematic empirical research on the diffusion of Internet technology to businesses.[2]

This paper fills this gap with both a novel framework and novel data. We estimate probit regressions of Internet adoption as a function of characteristics of a business establishment (i.e., its industry, location, and size and whether or not it is part of multi-

[1] See Bresnahan and Trajtenberg (1995) for formal development of the theory.

[2] To be sure, geographers and government agencies have studied the diffusion of Internet technology at length, but largely among households. For statistical research about business use of the Internet, see Atrostic, Gates, and Jarmin (2000), Mesenbourg (2001), Atrostic and Nyugen (2001), and Atrostic and Gates (2001), as well as the Census (2002), Varian et al. (2001), Whinston et al. (2001), Forman (2002), and Kraemer, Dedrick and Dunkle (2002). None have explored geographic variation in use.

establishment organization). Our analysis highlights why adoption decisions vary across locations.

There were many purposes for adopting the Internet in business. We contrast simple purposes with complex ones. The first purposes, together labeled *participation*, relate to activities such as email and web browsing. This represents the use of the Internet for basic communications. The second purposes, together labeled *enhancement*, relate to investment in frontier Internet technologies linked to computing facilities. These technologies are often known as "e-commerce," and involve complementary and complex changes to internal business computing processes.

Our study is a short-run analysis that holds establishment locations fixed. While it is too soon to observe the long-run movement of establishments in reaction to this diffusion, our findings inform research about the relationship between information technology and the geographic location of firms.[3] We highlight three key factors that induced variation in experience across locations and that support three theories that describe how these factors influenced Internet dispersion.

- **Overcoming agglomeration disadvantages:** Some researchers hypothesize that the Internet helped businesses bridge distances between geographically disparate economic actors. While all business establishments benefited from this increase in capabilities, geographically isolated businesses benefited comparatively more from their ability to transfer data and transact electronically in ways that were previously difficult or prohibitively expensive. We label this *global village theory*.
- **The user cost of the Internet:** Some researchers hypothesize that adoption, adaptation, and operations for the Internet were comparatively more expensive outside of urban areas. Internet infrastructure, access fees, maintenance and development were all

[3] See, e.g., Gasper and Glaeser 1998; Kolko 2002, Sinai and Waldfogel, 2000.

more expensive. We call this *urban density theory.* We further posit that complex applications of Internet technology are more sensitive to variation in geographic supply constraints than simple applications.

- **The geographic dispersion of IT-intensive businesses:** Establishments locate in places for a wide variety of reasons. Many of these decisions were made prior to the diffusion of the Internet. Some of these establishments belonged to industries that already made heavy use of information-intensive technologies, while others did not. Because many of these firms clustered together in the same cities, these urban locations were coincident with high demand for the applications associated with new technology. We call this *industry composition theory.* It produces heterogeneous response to newly available technology that is unrelated to (but nevertheless correlated with) location, and is essential for understanding whether new Internet technology becomes a substitute or complement to urban agglomeration.

Our econometric analysis performs several tests of GPT theory. First, we compare the relative strength of urban density theory and global village theory in both simple and complex applications, addressing a key debate in the literature on the diffusion of the Internet. That is, the contrast between simple and complex applications tests whether urban density theory is more relevant to a complex application than a simple one. By introducing industry composition theory into this debate, we demonstrate the precise sense in which agglomeration of industries in a city act as either a complement or substitute for new Internet technology.

This study examines Internet adoption at 86,879 establishments with over one hundred employees. Our sample is roughly half of all such establishments. This data comes from a survey updated to the end of 2000. Harte Hanks Market Intelligence, a commercial market research firm that tracks use of Internet technology in business, undertook the survey. We use the County Business Patterns data from the Census and routine statistical methods to generalize our estimates to the entire population of medium to large establishments in the United States. Approximately two thirds of the United States work force is employed in such establishments.

The results are consistent with GPT theory. First, participation is near saturation in a majority of industries and locations, which is evidence that benefits exceed costs for most establishments. It is also is consistent with global village theory. Once industry composition is controlled for, we find further weak support for global village theory; in other words, the results suggest that establishments in larger MSAs are less likely to adopt participation.

Second, the estimates for enhancement differ from participation in ways consistent with urban density theory. Controlling for industry composition, the percentage of establishments who adopt enhancement in large and medium MSAs is one full percentage point higher than in small MSAs and rural areas. This is a significant increase given that only 12.6% of firms undertake enhancement. The urban location of an establishment does contribute to observed outcomes.

Finally, we show why enhancement and urban areas were complements over a time period in which Internet technology was diffusing. We show that the relationship between geography and IT adoption will depend on the technology itself and on the period of study. Moreover, we show that the short-run adoption decisions of multi-establishment firms differ from those of single establishment firms. Consistent with previous research (e.g. Holmes 2002) and evidence from our data, we present a theory in which firms decide at which establishment to locate IT investment. Given constant benefits across an organization, multi-establishment firms are more likely to adopt IT where costs are lowest: in urban areas.

We also offer a very different explanation from the prevailing literature on the digital divide. Mostly focusing on what we call participation, some authors argue that geographic usage patterns provide evidence of the existence of a digital divide (i.e., that the economic consequences of the Internet exacerbate regional inequalities, concentrating benefits in only a few locations).[4] Our data show that this characterization is misleading for business use of the Internet. Business participation, if anything, supports a more positive assessment. Moreover, business use of the Internet for enhancement is shaped by the prior geographic distribution of industry. Since many industries are not concentrated in a few locations, the Internet's advanced uses did not concentrate in a small number of areas. Instead, it dispersed to the vast majority of areas in the United States.

2. Background and Framework

Why are GPTs adopted at different rates by different firms? Below we offer a standard framework that describes how the benefits and costs of adoption can vary across firms.[5] We develop links between this framework and a number of theories that explain how the rate of Internet diffusion may vary by geographic location.

2.1 Theories of Internet Diffusion

Consider an establishment's decision to adopt a new technology such as Internet access. Below we will consider both a simple and a complex application of the Internet to business.[6] Establishment i will adopt Internet technology by time t if

---

[4] Articulation of this view is particular prominent in the National Telecommunications Information Administration's "Falling Through The Net" series (NTIA 1995, 1998, 1999, 2000). See, also, Moss and Townsend (1997), or Zooks (2000a, b).
[5] For an elaboration of this framework in the context of the diffusion of IT generally and heterogeneous experience across location, see Bresnahan and Greenstein, 2001.
[6] Our data analysis we will consider adoption of two distinct layers of Internet technology. The analysis above applies to both layers of Internet investment, so for notational simplicity we consider only a generic application of technology in this section.

$$NB(x_i,t) \equiv B(x_i,t)-C(x_i,t)>0$$

Where *NB* is the net benefit of adoption, *B* is the (gross) benefit of adoption, and *C* is the cost of adoption.[7] We define x_i to be a vector of attributes distributed across the population of establishments. For example, x_i may describe variation in geographic conditions, industry, or prior investments.

We observe establishments after they have made their choices. In other words, we observe whether an establishment made a decision to adopt or not adopt under a given set of conditions. We propose three contrasting theories on how location can influence the decision to adopt by time *t.* In these theories, x_i represents variables indicating the size of urban area and the density of population.

- **Global village theory** argues that adoption benefits are decreasing as urban density and size increases (i.e., $dB/dx_i < 0$, where x_i is density). Firms in areas with low population or low population density will benefit more from new Internet technology. This view has received considerable exposure (e.g., Cairncross 1997, Castells, 2001) but little empirical verification.[8] In its most basic form, global village theory argues that new Internet technology helps break down communication barriers between individuals and organizations, and these barriers are greatest for geographically isolated firms. That is, establishments in rural or small urban areas will derive the most benefit from Internet technology because their suppliers and customers are most likely to be located in other areas. Moreover, these areas will derive higher benefits because of a lack of substitute

[7] Generally, see Rogers (1995). Because our data are cross-sectional, we examine the decision for an establishment to adopt by a certain date *t*. We allow the cost term *C* to include the opportunity cost of not adopting at some other time $s > t$, thus the profitability condition above is both necessary and sufficient for the establishment to adopt by *t.* A more traditional formulation would examine an establishment's decision to adopt at time *t*, and the equation above would be supplemented by an "arbitrage condition" (Ireland and Stoneman 1986) that it is unprofitable to adopt at any other time $s \neq t$.

[8] One exception is Forman (2002), who examines a sample of service sector firms, and shows that rural firms and firms that are geographically dispersed are more likely to adopt Internet technology. Also see Premkumar (2000) and Premkumar and Roberts (1999) for a close study of rural business Internet adoption and use.

data communication technologies, such as fixed private lines. In other words, advanced Internet technology can substitute for many of the disadvantages associated with a location remote from the urban center of economic activity. Internet technology is a substitute for agglomeration.

- **Urban density theory** argues that adoption costs are decreasing as population size and density increases (i.e., $dC/dx_i < 0$, where x_i is density). There are three major reasons for this hypothesis: (1) availability of complementary information technology infrastructure, (2) labor market thickness, and (3) knowledge spillovers. These are closely related to the three major reasons given for industrial agglomeration (e.g., Marshall 1920; Krugman 1991). The availability of low-cost complementary inputs, such as broadband services or Internet access, will have a major impact on the cost of adopting participation or enhancement.[9] Labor market thickness is closely related to complementary inputs. Adoption of new technologies often requires specialized technical skills not available within the firm and which must be procured either by hiring additional workers or through arms-length contracting. In other words, locating in an urban center acts as a complement to the use of advanced Internet technology. In shorthand, Internet technology is a complement to urban agglomeration.

- **Industry composition theory** asserts that demand for the Internet increases with urban density because information-intensive firms tend to inhabit more urban areas (i.e., $dB/dx_i > 0$, where x_i is density). Industry composition theory relies on two premises. First, some establishments place higher value on information-intensive activities than others in ways that vary systematically across industry and establishment size. Second, establishments from the same industry tend to cluster in similar places for many reasons

[9] Although,, by this time period, almost all geographic areas were serviced by Internet Service Providers (Downes and Greenstein, 2002). There were, however, some exceptions in especially poor or remote areas. See Strover, Oden, and Inagaki(2002).

to take advantage of thicker labor markets and other shared local resources (Krugman, 1991). Concentration of Internet technology-intensive activity in one location could have little to do with location-specific benefits. If previous decisions to concentrate activity results in the clustering of some types of firms in urban areas, then it can result in a concentration of adoption of new Internet technology in urban areas.

2.2 The Predictions of Different Theories

We highlight five predictions of global village, urban density and industry composition theories, while taking into account the simple and complex categories of adoption, as well as whether the firms adopting the Internet were single- or multi-establishment businesses:

- **Global village theory versus urban density theory:** If the location of a business establishment is fixed prior to the diffusion of technology, and unchanging throughout its diffusion, then these two theories give opposing predictions about the relationship between adoption and urban density. Controlling for establishment industry, if we observe adoption increasing with density, then we infer that this supports urban density theory over global village theory. If we observe adoption decreasing with density, then we infer the opposite.
- **Simple versus complex applications of a GPT:** Studies of the adoption of computing technology argue that co-invention costs shape investments by business users (Bresnahan and Greenstein 1996). Co-invention expenses arise when a GPT is adapted for commercial use. Co-invention requires high-quality labor inputs or third-party technological mediators that may cost more in sparsely populated areas than in urban areas. In the case of the Internet, co-invention involved the local demands of business establishments and the supply constraints of markets for Internet technology infrastructure and services. Previous work (e.g., Forman 2002) have shown that some

simple applications were viable as soon as the Internet commercialized, while other complex applications required time, expense and invention to become viable with business users.[10] Therefore, co-invention theory forecasts that urban density theory is likely to be stronger for the diffusion of a complex application. That is, $dC_C/dx_i < dC_S/dx_i < 0$, where C_C is the cost for a complex application and where C_S is the cost for a simple application. In words, the combination of urban density and co-invention theories suggest that adoption of a complex technology should be more sensitive to increases in population density than adoption of a simple one.

- **Urban density theory versus industry composition theory:** When we observe only adoption and location, these two theories make observationally equivalent predictions about the sensitivity of net benefits to urban density. Both predict $dNB/dx_i > 0$, that is, the net benefits of adoption increase with urban density. Industry composition theory predicts additionally that this relationship is due to pre-existing industry composition rather than direct benefits of urban agglomeration. Given the industry, industry composition theory predicts $dNB/dx_i=0$.
- **Cities as complements for IT:** It is possible that urban density theory and industry composition theory interact with each other so that IT-friendly cities get a larger benefit from industries that are IT-friendly than do other cities. In particular, IT-intensive industries may gain extra benefits due to spillovers from being in an area with a high adoption rate. This story should be particularly important for enhancement as learning about frontier technologies may require pooled resources.

The opposite result is also possible. IT-intensive industries may have much expertise in-house and have less need to take advantage of the thicker labor markets

[10] Greenstein (2001) provides such an argument for the development of the Internet access industry in the United States. Internet access was comparatively simple. Use of the Internet for e-commerce typically was not. See Forman (2002) for a discussion of the latter during the early diffusion of the Internet.

associated with IT-friendly locations. This labor market thickness story should matter much more in areas with low populations because this story relies on absolute labor market size more than on percentages.

Formally, if cities and IT are complements, $dNB^2/dIdx_i > 0$ where I is the industry's tendency toward IT adoption and x_i is density. If cities and IT are substitutes $dNB^2/dIdx_i < 0$. We expect $dNB^2/dIdx_i > 0$ for larger cities and enhancement and $dNB^2/dIdx_i < 0$ for smaller cities and participation.

- **Multi-establishment firms and urban density:** The benefits and costs to adopting Internet technology may vary depending on whether the establishment is part of a multi-establishment firm. Multi-establishment firms often adopt new communication technologies at some, but not all, of their establishments. In establishment-level analysis, this will have no effect on global village, $dB_M/dx_i = dB_S/dx_i > 0,$ where B_M is the benefit for an establishment in a multi-establishment firm and where B_S is the benefit for an establishment in a single establishment firm. In other words, the benefits of geographic isolation are independent of whether the establishment is part of a multi-establishment firm. However, multi-establishment firms will choose to locate Internet technology in the least-cost locations, implying that the effects of greater population size and density (urban density) will be more important to adoption decisions for establishments part of multi-establishments firms. This will be particularly true for complex enhancement technologies, where co-invention costs are higher.

3. Data and Method

The data we use for this study come from the Harte Hanks Market Intelligence CI Technology database (hereafter CI database).[11] The CI database contains establishment-level data on (1) establishment characteristics, such as number of employees, industry and location; (2) use of technology hardware and software, such as computers, networking equipment, printers and other office equipment; and (3) use of Internet applications and other networking services. Harte Hanks Market Intelligence (hereafter HH) collects this information to resell as a tool for the marketing divisions at technology companies. Interview teams survey establishments throughout the calendar year; our sample contains the most current information as of December 2000.

Harte Hanks tracks over 300,000 establishments in the United States. Since we focus on commercial Internet use, we exclude government establishments, military establishments and nonprofit establishments, mostly in higher education. Our sample contains all commercial establishments from the CI database that contain over 100 employees, 115,671 establishments in all; [12] and HH provides one observation per establishment. We will use the 86,879 observations with complete data generated between June 1998 and December 2000. We adopt a strategy of utilizing as many observations as possible because we need many observations for thinly populated areas.[13] This necessitates

[11] This section provides an overview of our methodology. For a more detailed discussion, see Forman, Goldfarb, and Greenstein (2002).

[12] Previous studies (Charles, Ives, and Leduc 2002; Census 2002) have shown that Internet participation varies with business size and that very small establishments rarely make Internet investments for enhancement. Thus, our sampling methodology enables us to track the relevant margin in investments for enhancement, while our participation estimates may overstate participation relative to the population of all business establishments.

[13] If we were only interested in the features of the most populated regions of the country, then we could easily rely solely on the most recent data from the latter half of 2000, about 40% of the data. However, using only this data would result in a very small number of observations for most regions with under one million in population.

routine adjustments of the data for the timing and type of the survey given by HH. Table A1 in the Appendix compares the HH data with the Census data.

3.1. Sample Construction and Statistical Method

Using two survey forms, HH surveyed establishments at different times. To adjust for differences in survey time and type, we econometrically estimate the relationship between an establishment's decision to participate or enhance as a function of its industry, location, timing of survey and form of survey.

To be precise, our endogenous variable will be y_j. The variable y_j is latent. We observe whether or not the establishment makes a discrete choice, for example, chooses either participation or enhancement. In either case, the observed decision takes on a value of either one or zero. We will define these endogenous variables more precisely below.

We assume that the value to an establishment j of participating in the Internet is

$$(2) \qquad y_j = \sum_i \alpha_i d_{ij} + \sum_l \beta_l d_{lj} + \sum_t \gamma_t d_{tj} + \sum_{t>19905} \delta_t d_{tj} d_{pj} + \sum_p \phi x_{pj} + \varepsilon_j \ ,$$

where d_{ij} and d_{lj} are dummy variables indicating the industry and location of the establishment, d_{tj} indicates the month in which the establishment was surveyed, and d_{pj} indicates whether the establishment responded to the long survey.[14] The variables x_{pj} denote other location-specific and establishment-specific variables, such as population size, density, firm size, and whether the firm is a single- or multi-establishment firm, as well as interactions between these variables. If we assume the error term ε_j is i.i.d. normal, then

[14] HH used two surveys. One asked for more details on IT use than the other. We interact the long survey dummy with time. This controls for endogeniety of survey. If establishments are selected for the long survey endogenously by HH, then the impact of receiving the long survey on adoption may vary over time.

the probability that establishment j participates is a probit. The probability of adopting enhancement can be written similarly.

We use this model for two research purposes. Our first purpose is descriptive. We illustrate average tendencies by predicting the average adoption probabilities for particular locations at a particular point in time. We then weight observations using Census County Business Patterns data to obtain a representative sample. We do this to inform the reader about the basic patterns of the endogenous variable. For the average estimates in Tables 1, 2, 3, we calculate predicted probabilities of adoption for each establishment *as if it were surveyed in the second half of 2000 and were given the long survey*. The x_{pj} are not included in this specification. Once we weight by the true frequency of establishments in the population, we have information about establishments related to two-thirds of the United States workforce. This exercise will provide the first comprehensive census of commercial Internet use.

Our second purpose is investigative. We analyze the marginal contribution of different factors that shape the adoption decision at the establishment. We report marginal effects from a variety of different specifications. These are the coefficients on ϕ, weighted to give a representative sample. Tables 4, 5, 6, and 7 and Figures 1 through 4 display these marginal effects. Three econometric assumptions support the estimates of marginal effects:

- **Exogenous location:** We examine short-run marginal effects of industry and location variables on the decision to invest in Internet technology. To identify these effects, we assume that the location of an establishment is exogenous. This assumption is both plausible and testable. First, it is plausible because we examine large establishments, where the relocation costs are highest, and where, as a practical matter, relocation churn in the US economy is the lowest. Moreover, the commercial Internet surprised most of these firms in 1995, so it is implausible that most of them located in a specific place in anticipation of the Internet's diffusion. Second, we can test this assumption directly by comparing results between our entire sample of establishments and a special sub-sample of establishments who (we are certain) fixed their locations prior to the availability of the commercial Internet, i.e., prior to 1995. If the key estimates do not differ between these two samples, we infer that the potential endogeneity of establishments does not alter our inferences about the influence of location on adoption of Internet technology.

- **Simultaneity bias:** Our base econometric specification assumes that the adoption decision of one establishment is independent of every other. This assumption is potentially questionable for multi-establishment firms in which a central executive decision maker (e.g., a CIO) possibly coordinates the choice to adopt or not adopt for all establishments. Depending on a wide variety of factors, in theory adoption decisions at establishments from the same organization could be either substitutes or complements for one other. While understanding that relationship is of independent interest, it also lies outside the scope of this study. For purposes of this study, we are concerned that simultaneity influences the coefficient of interest, the estimate of location on adoption at each establishment. We address these concerns directly by characterizing the decisions of related establishments at other locations in a reduced form way, then measuring whether this alters the estimate of the coefficient on location, instrumenting for decisions elsewhere. Once again, our focus will be on whether our inferences about the influence of location on adoption of Internet technology are robust to introducing simultaneity into the estimation.

4. Defining Simple and Complex Applications

As a GPT, Internet technology is employed in many different uses and applications. Our sample includes at least twenty different types of IT, from basic access to software for TCP/IP-based Enterprise Resource Planning (ERP). Moreover, there are considerable differences in the applications used across establishments, rendering an application-based measure of adoption inadequate for our purposes. Instead, we identify two types of applications: *participation* and *enhancement*.[15] The first is simple and requires little co-invention. The second is complex and requires significant co-invention.

Participation: Participation is affiliated with basic communications, such as email use, browsing, and passive document sharing. It represents our measure of the minimal Internet investment required to do business on the Internet. It is emphasized in many studies of ubiquitous communications networks. A ubiquitous network is one in which every potential participant is, in fact, an actual participant. Concerns about ubiquity emerge in policy debates about applying principles of "universal service" to new technologies (Cherry, Hammond and Wildman 1999, Compaine 2001, Noll et al. 2001). Geographic differences in adoption, such as urban/rural divisions, are important drivers of policy decisions in this area.[16]

[15] An alternative strategy would be to treat Internet technology as fungible IT capital and employ an adoption measure that treated all Internet investment equally. Such a strategy would ignore the considerable heterogeneity in costs and benefits across purposes.

[16] To be counted as participating in the Internet, an establishment must engage in two or more of the following activities: (1) have an Internet service provider; (2) indicate it has basic access; (3) use commerce, customer service, education, extranet, homepage, publications, purchasing or technical support; (4) use the Internet for usage, or has an intranet or email based on TCP/IP protocols; (5) indicate there are Internet users or Internet developers on site; or (6) outsource some Internet activities. We looked for two or more activities to guard against "false positives". As it was, this was a minor issue. Most respondents responded affirmatively to many of these criteria.

Enhancement: Enhancement, on the other hand, is affiliated with IT that either changes existing internal operations or implements new services.[17] Enhancement is linked to the productive advance of firms and the economic growth of the regions in which these firms reside. It usually arrives as part of other intermediate goods, such as software, computing or networking equipment. Benefits accrue to the establishment that employs enhancement through the addition of competitive advantage, but the costs and delays of this activity vary widely.[18]

Identifying participation was simple compared to identifying enhancement. We identify participation when an establishment has basic Internet access or has made any type of frontier investment. The establishment survey gives plenty of information about these activities, so we identify participation with confidence. In contrast, enhancement activity is less transparent in the survey. We look for indications that an establishment must have made the type of investment commonly described in books on e-commerce. We identify enhancement from the presence of substantial investments in e-commerce or e-business applications. The threshold for "substantial" is necessarily arbitrary within a range. To provide confidence that we are measuring substantial investment, we look for commitment to two or more of the following projects: Internet-based enterprise resource planning or TCP/IP-based applications in customer service, education, extranet, publications, purchasing or technical support.[19] [20]

[17] See for example, Malone, Yates, and Benjamin (1987), Hubbard (2000), Hitt and Brynjolfsson (1997), or Bresnahan, Brynjolfsson, and Hitt (2002).
[18] This varies from the definitions employed by Porter (2001). This is due to a difference in research goals. Throughout his article, Porter discusses the determinants of, and shifting boundaries between, investments that provided table stakes and those that complement a firm's strategy and enhance competitive advantage. He argues that these levels vary by industry and differ from firm to firm. This is the proper variance to emphasize when advising managers about their firm's strategic investment. In contrast, our measurement goals require both a standardized definition and a consistent application across industries and locations.
[19] We tested slight variations on this threshold and did not find qualitatively different results.
[20] For a more precise definition describing exactly how enhancement was coded, see Forman, Goldfarb, and Greenstein (2002).

5. The Dispersion of Participation and Enhancement

In this section, we seek to characterize differences in average participation and enhancement rates across industries and locations. To obtain a representative sample, we compared the number of firms in our database to the number of firms in the Census. We calculated the total number of firms with more than 50 employees in the Census Bureau's 1999 County Business Patterns data and the number of firms in our database for each two-digit NAICS code in each location.[21] We then calculated the total number in each location. This provides the basis for our weighting. The weight for a given NAICS in a given location is

$$(1)\quad \frac{Total\ \#\ of\ census\ establishments\ in\ location-NAICS}{Total\ \#\ of\ census\ establishments\ in\ location}\cdot\frac{Total\ \#\ of\ establishments\ in\ our\ data\ in\ location}{Total\ \#\ of\ establishments\ in\ our\ data\ in\ location-NAICS}.$$

Each location-NAICS is given its weighting from its actual frequency in the Census. In other words, if our data undersamples a given two-digit NAICS at a location relative to the Census then each observation in that NAICS-location is given more importance.

In Table 1 we present average rates for participation and enhancement for the US. Participation by establishments within the sample is at 80.7% (see Unweighted Average in Table 1). The sample under-represents adopters. Our estimate of the economy-wide distribution, using the true distribution of establishments from the Census, is 88.6% (see

[21] We use 50 employees because potential differences between different times for taking the survey mean that firms could grow after the Census and therefore be in the CI database. It was necessary to be inclusive for the weighting because some small rural areas had less than three firms in both the Census and the CI database; and therefore if one firm grew from the time of the Census to the time of the CI survey, the weightings would be difficult to interpret. The results are robust to weighting by firms with more than 100 employees in the Census and those with more than 25 employees. This is not surprising given the high correlation between these values.

Weighted Average in Table 1). Enhancement has been undertaken by 11.2% of our sample and 12.6% of the true distribution.

We also can estimate the rate of enhancement adoption by "experimenters," that is, by those establishments with *some* indication of enhancement use, but not much. As one would expect for a technology still in the midst of diffusion, the proportion for experimenters (combined with enhancement) is considerably higher than for enhancement alone, reaching 18.1% for the unweighted average and 23.2% for the weighted average. We have explored this latter definition and found that it tracks the enhancement definition we use below, so it provides no additional insight about the dispersion of use. We do not analyze it further.

In Tables 2a and 2b we provide an overview of participation and enhancement adoption results for the largest economic areas in the United States. We list the estimates organized by Metropolitan Statistical Areas (MSAs), with over one million people and in order of highest to lowest adoption rates.[22] As we do in all of our tables, we list the standard errors and number of observations to identify the degree of statistical confidence in the estimates.[23] (For comparison, Tables 2a and 2b also list the marginal effect of location on adoption, which we will discuss later.)

In Table 2a, we show that participation is high in major urban locations. Virtually all establishments in the major urban areas are participating. We estimate that of the forty-nine MSAs, thirty-five are above 90%. All but five are within a 95% confidence interval of 90%. Big differences among metropolitan areas are apparent only at the extreme. The bottom ten areas range from 89.1% in Pittsburgh to 84.6% in Nashville. Although these are

[22] When two or more MSAs are part of the same urban environment, the census combines them into CMSAs. For example the Dallas-Fort Worth CMSA contains both Dallas and Forth Worth. In Table 2 we present the CMSA results rather than the individual MSA results when an MSA is part of a CMSA.

[23] These are computed using the delta method.

the lower adopting areas, they are not very low in absolute value. Participation is nearly universal in large cities.

In Table 2b we examine the use of enhancement at establishments in MSAs with over one million people. All but one are within a 95% confidence interval of the national average of 12.6%. The top ten include a set of areas that partially overlaps with the list in Table 2a. (Five of the top ten are also in the top ten for participation.) The list begins with the greater Denver area (with 18.3%) at number one and the greater Portland area at number ten (with 15.1%). In between are the greater San Francisco Bay Area, the greater Salt Lake City area, Minneapolis/St Paul, the greater Houston area, Atlanta, Oklahoma City, Dallas/Fort Worth, and San Antonio. The bottom ten areas range from 12.4% in Phoenix to 9.0% in Las Vegas. Even so, these low adopting areas are, once again, not very low relative to the average.

In Table 3 we further examine general tendencies by showing participation and enhancement rates across types of locations in the United States. In contrast to Tables 2a and 2b, Table 3 shows sizable differences in participation and enhancement between large urban, small urban, and rural areas. On the surface, this evidence supports either urban density theory or industry composition theory. We see that large MSAs are somewhat exceptional, with an average participation rate of 90.4%. Participation rates in medium-sized MSAs and rural (non-MSA) areas are lower at 84.9% and 85.1%, respectively. In small MSAs the participation rates are even lower, 75.5% on average.[24]

The disparities in adoption rates for enhancement are even greater (see Table 3). Again large MSAs are somewhat exceptional, with average adoption rates of 14.7%. In medium MSAs, adoption averages 11.2%. In small MSAs the rates are even lower, 9.9%

[24] From this point forward, MSAs with populations greater than 1 million will be referred to as *large MSAs*, those with between 250,000 and 999,999 will be *medium MSAs,* those with less than 250,000 will be *small MSAs*, and non-MSA areas will be called *rural*.

on average. Average adoption rates in large MSAs are almost one-third greater than in medium MSAs. Once again, these averages suggest that the urban density theory or industry composition theory may hold. To identify between them, we will need to identify the marginal effect of location while holding constant industry composition. We address this in the next section.

6. The Marginal Impact of Population Concentration

6.1 The Marginal Impact of Location

In this section, we continue to estimate equation (2), but our focus is now on estimating the parameters ϕ. We weight observations by the inverse probability that an establishment will appear in our sample. To be precise, the weight for each observation is the total number of establishments in a state/NAICS in Census County Business Patterns data divided by the number of establishments in the state/NAICS in our sample multiplied by controls for sampling the same establishment twice.

Table 4a shows the results of probit regressions, and Table 4b presents the marginal effects. All probit regressions include dummy variables for 3-digit NAICS, the month the data were collected, survey type, survey type interacted with month, and whether or not the establishment was part of a multi-establishment firm. Employment and employment squared were also included as controls. Population was measured at the MSA level and density at the county level.

For columns 1 and 5, we use nonurban (hereafter termed *rural*) state areas for the base. For columns 2, 3, 6, and 7, we include a "rural area" dummy for rural areas, since no meaningful population figures exist for these areas. In columns 4 and 8 we include population density for all urban and rural areas using low-density areas as the base.

Participation: From Table 4, it is clear there is no support for the urban density theory. Controlling for industry and firm characteristics, location size and density has little impact on the decision to adopt at the participation level. If anything, the effects of location size and density support the global village theory. To be sure, the impact of geography is of limited economic and statistical significance, so the support is weak. For column 1, we use nonurban state areas (i.e., non-MSAs) for the base, and the results in Table 4b show that medium and large MSAs are 0.5% to 1.0% less likely to have adopted participation by the end of 2000. However, the effect is only significantly different from rural areas for medium MSAs. Moreover, this effect is only of marginal economic significance as participation rates average 88.6%.

In column 2, we identify the effects of size through a variable that captures the effects of increases in population in urban areas. Increases in population size decrease the probability of participation, though the effects are statistically insignificant. In columns 3 and 4, we explore further specifications. Including a squared population term implies that the effects of population will turn negative once urban areas exceed 7.039 million, a threshold that is larger than all but the five largest urban areas. In column 4 we include dummies for population density. Density is measured at the county level and is split into quartiles, with the dummy for the bottom quartile omitted. This alternative specification gives very similar results. Variation in population density does not affect participation by more than 1%, and is always statistically insignificant.

Enhancement: The effects of population size and density on enhancement support the urban density theory. Column 5 in Table 4b shows that establishments in medium and large MSAs adopt enhancement at a rate of 0.8% to 1% higher than small MSAs and rural areas. Column 8 shows that establishments in medium- and high-density regions adopt enhancement at a rate of 1% to 1.5% more. All of these effects are statistically significant.

They are economically significant in light of the average enhancement rates of 12.6%. While column 6 suggests that a linear population term has little effect on enhancement, column 7 shows that population will have a statistically and economically significant positive effect for all metropolitan areas below 8.8 million in size (all but New York, Los Angeles, and Chicago).

The contrast between participation and enhancement is informative. As GPT theory predicts, the support for urban density theory is stronger for the more complex applications. In Figures 1 and 2 we graph the marginal effect of location in the baseline probit in model (1).[25] We divide locations into four types: large MSAs, medium MSAs, small MSAs, and statewide rural (non-MSA) regions. In Figures 1 and 2 we plot the kernel density estimates of the effects of location on participation and enhancement, respectively.[26] We use Epanachnikov kernels with "optimal" bandwidths.

In Figure 1 small MSAs and rural areas have a fatter right tail, while the density for large MSAs reaches its peak below any of the three other classes of geographic areas. Table 5 shows that large and medium MSAs have lower average and median marginal effects than rural areas or small MSAs, however the differences in average marginal participation effects are not statistically significant. In all, these figures provide further support for global village theory: increases in local population size and density do not increase the likelihood of participation adoption. If anything, they lower it.

In Figure 2, it is clear that the density estimate for large MSAs stochastically dominates those for small and medium MSAs and rural areas. Table 5 shows that the average and median marginal effects for enhancement are larger (more positive) than those for any of the other classes. However, the average marginal enhancement effects for large

[25] This probit is not depicted in any table. We identify the effects of population size and density directly through the marginal effects.
[26] The omitted MSA is San Jose, the top MSA in adoption of enhancement.

MSAs are very close to those for small MSAs (-0.00652 versus –0.00708). The median enhancement marginal effect for medium MSAs is larger than rural or small MSAs; however, the average is lower than that for small MSAs. In all, the statistical evidence shows that the very largest MSAs are more likely to adopt enhancement than establishments in other types of geographic locations, which provides support for the urban density theory for enhancement technology.

Robustness: We checked our results for a variety of robustness issues. As noted, we were concerned that establishment location decisions might be endogenous with improvements in communications technology. We re-estimated the model using only establishments that had been added to the HH database prior to 1995, the year in which Internet technology began to diffuse widely to businesses. Although this restricted the size of our sample substantially (to 23,436 observations), the basic results remain the same. The correlation coefficients between our baseline marginal effects and those using pre-1995 data are 0.829 for participation and 0.997 for enhancement.

We tried a number of other robustness checks. For one, we worried about omitted variables. We experimented with a variety of different specifications by using different location variables (e.g., CMSA dummies), different firm controls (e.g., revenue, private/public), and alternative measures of population size and density. We also worried about how weighting the probit model would affect our results. We tried weighting the probit regressions by 3-digit NAICS/states, 2-digit NAICS/MSAs, as well as trying no weighting at all. In all cases the results remained qualitatively the same; the correlation coefficients between our baseline coefficient estimates and the alternative specifications were between 0.88 and 0.95 for participation and 0.78 and 0.90 for enhancement. Furthermore, qualitative results did not change.[27]

We also checked the robustness of these estimates to different specification of the adoption decision at establishments that are part of a multi-establishment organization. Our key results are largely robust to these specifications, though there is some complexity to these findings. For the sake of narrative continuity, we defer discussion of this complexity momentarily, reporting it in detail below.

6.2 Competing Marginal Effects: Industry v. Location

Industry composition: The differences between Tables 3 and 4 show that the effects of location on participation and enhancement fall if controls for establishment size, industry, and firm status are included. In Table 3, large MSAs have almost a 15% higher participation rate and 5% higher enhancement rate than small MSAs, whereas in Table 4, locating in a large MSA rather than small MSAs reduces the probability of participation by 0.8% and increases the probability of enhancement by 0.8%. This supports industry composition theory.

The large differences in adoption rates between large and small urban areas in Table 3 reflect differences in industry composition across locations. Industry composition explains much more of the variation in participation and enhancement rates than location. Once industry is controlled for, the incremental contribution of location in the probit regressions is small. This is shown in Table 6. The pseudo-R^2 of a probit for participation including location dummies only is 0.1526, whereas the pseudo-R^2 of a probit with

[27] We also explored whether systematic establishment differences across geographic locations are driving our results. Though unobservable establishment differences could play a role, we were unable to uncover any pronounced observable establishment differences. Using weighted data, large MSAs are larger (12.8% of establishments with >500 employees versus 9.0% for small MSAs and 9.9% for rural areas) and more likely to be multi-establishment (48.7% multi-establishment versus 43.9% for small MSAs and 33.4% for non-MSAs). However, when using unweighted data much of the differences disappear (12.7% of establishments in large MSAs >500 employees versus 12.1% in small MSAs 13.4% in rural areas; 46.3% of establishments in large MSAs multi-establishment versus 46.8% in small MSAs and 40.8% in rural areas).

industry dummies only is 0.2251. Adding location dummies to a probit that includes industry dummies improves the fit only marginally, from 0.2251 to 0.2339.

Enhancement displays a similar pattern. Location dummies explain only 0.0347 of the variation in enhancement, industry dummies explain 0.0591, and the combination of industry and location dummies explains 0.0672. While there remains a great deal of unexplained variation in our results, it is clear that an establishment's industry explains more than geographic location.

***The Marginal Effects of Industry Composition*:** Our previous results suggest that industry composition explains more of the variation in participation and enhancement rates than geographic location. For industry composition theory to be true, however, leading industries must also be concentrated in large urban areas. To test this hypothesis, we separated establishments by geographic location (e.g., rural, small, medium, large MSA) and calculated the kernel density for each type of location. For each distribution, the underlying marginal effects are the same, however the densities of each marginal effect differ. We did this both for participation and enhancement.

In Figure 3, we show the kernel density estimates of the marginal effects of industry by geographic area for participation.[28] Lead-user industries tend to be concentrated in large geographic areas. The average of the marginal effects of industry in rural and small MSAs is –18.7% and –20.2%, while the marginal effects for medium and large MSAs are –18.8% and –16.9%. These averages are statistically significant from one another at the 1% level. Large MSAs tend to have more lead-user industries, even for participation.

[28] All industry results are unweighted. The omitted industry is information and data processing (NAICS 514). We use Epanachnikov kernels with bandwidth of 0.05 for participation and 0.005 for enhancement. These are wider than "optimal" bandwidths. Optimal bandwidths fail in this case as there are thousands of observations but the only 81 possible values that are there are the 81 relevant 3-digit NAICS levels. Therefore the optimal bandwidth does almost no smoothing.

Figure 4 shows that lead users of enhancement are even more skewed toward large MSAs. Rural areas and small MSAs have the highest densities along the left tail of the distribution, whereas large and medium MSAs dominate the right tail. The average marginal effect of industry on enhancement adoption is increasing in location size: –8.0% in rural, –7.8% in small MSAs, –7.7% in medium MSAs, and –7.4% in large MSAs. Again, these are significant at the 1% level. The results in Table 3 reflect differences in industry composition between small and large MSAs, rather than other location-specific benefits of locating in urban areas.

Figure 4 provides foundation for an intuitively appealing observation—urban areas are comprised of establishments with a disproportionate tendency to be information intensive. To be concrete, within large MSAs, 27.5% of establishments are in industries that are part of the top quartile of adopters, compared to 19.0% of establishments in small urban areas. The industries in the upper quartile are traditionally information intensive, such as company headquarters, utilities, finance and insurance, professional and scientific services, electronics manufacturing, and wholesale trade.[29] The geographic dispersion of establishments from these industries favored large urban areas prior to the diffusion of the Internet and largely contributed to higher rates of participation and enhancement in large urban areas. We further illustrate these results in Tables 2a and 2b, which rank large MSAs by their average rates of participation and enhancement. The third column includes the rank of location, after controlling for industry and other features of the establishment.

Average participation rates across business establishments combine effects from both a location and industry composition. The Spearman correlation between the average and marginal rankings is 0.5389. The San Francisco Bay Area achieves a high average participation by having industries with high participation (e.g., electronics manufacturing)

[29] For more detail, see Forman, Goldfarb and Greenstein, 2002.

and a favorable marginal location rank (seventh). In contrast, Denver achieves high participation in spite of low marginal location rank (forty-third) and due to the composition of establishments from industries with high participation (e.g., electronics manufacturing). Interestingly, the pattern does not hold at the lowest end. The areas with lower average participation have both low marginal rankings for their areas and the compositions of their industries do not make up for those rankings.

Table 2b shows that the marginal contribution of location is mildly more important. The Spearman correlation between the average and marginal ranking is 0.711. In the top ten, all but one (i.e., Atlanta) of the areas with high average adoption also have high marginal ranking for their location. In the bottom ten, all but one (i.e., Sacramento) of the areas with low average adoption also have low marginal rankings for their location.

Complementarities between industry and location**:** The marginal contribution of industry may differ by location; however the direction of this difference is ambiguous. IT-intensive industries may gain extra benefits due to spillovers from being in an area with a high adoption rate. On the other hand, these industries may have much expertise in-house and have less need to take advantage of the thicker labor markets associated with IT-friendly locations. We test for this relationship by comparing median industries by IT use across cities.

For participation, the marginal contribution of the median industry in a location is uncorrelated with the marginal contribution of the location in general (ρ=-0.0214). However, this result disguises a large difference between large and medium MSAs on the one hand and small MSAs and rural areas on the other. The correlation between the marginal contribution of median industry and location is significantly positive (ρ=0.307) for large and medium MSAs but significantly negative (ρ=-0.211) for small MSAs and rural areas. For larger population locations, good cities do have good industries for participation; yet for smaller areas the opposite is true.

For enhancement, there do seem to be complementarities between cities and locations (ρ=0.161). These complementarities do not vary much by city size. Regardless of size, good industries are in good cities and good cities are dominated by good industries. Cities and IT appear to be complements, except in the case of participation in low population areas. The complementarity is likely a result of spillover effects in using frontier technologies. The exception is as interesting as the more general finding. In low population areas, favorable locations only help adoption in firms without in-house expertise. In that case, location and industry become substitutes.

Consistency with GPT Theory: Overall, the average rate of participation is quite high, almost 90% in this population (i.e., in large MSAs) of establishments. This is consistent with the view that participation in the Internet was simple and had low user cost, which supported participation adoption over a wide range of potentially low or high benefits. Many establishments adopted in spite of their location because costs were already low. Urban density or other local factors did not make matters markedly better or worse.

Using the Internet for enhancement, however, involved much higher co-invention costs. The adoption and development of more expensive and more complex purposes involved greater costs. Those costs did rise and fall with location, so population density shaped decision-making on the margin. An establishment could lower or raise the cost of enhancing its computing facilities with the Internet if it had the good or bad fortune of being located in a favorable area. This result continues to hold even when controlling for industry composition.

6.3 Internet Adoption in Multi-Establishment Firms

As noted above, the marginal effect of location on adoption will vary depending on whether an establishment is from a multi-establishment firm. In this section, we systematically examine how multi-establishment status changes the returns to location. To our knowledge, this study is the first to comprehensively address how multi-establishment status affects the returns to technology adoption.

Average participation and enhancement rates differ significantly across single-establishment and multi-establishment firms. Participation is adopted at an average of 85.9% of establishments among multi-establishment firms, whereas the comparable statistic for enhancement is 23.3%. That is, average unconditional participation is mildly lower than participation in single establishment firms, while average unconditional enhancement is much higher. If establishments are aggregated up to their parent organization, our data show that participation and enhancement happens for at least one establishment in, respectively, 99.4% and 59.5% of multi-establishment organizations in our sample.[30] In this respect, average unconditional participation of an organization is higher than participation for single establishment organizations, and so too is average unconditional enhancement.

It is, of course, more complicated to predict adoption of Internet technology at the establishment level, conditional on belonging to a multi-establishment organization. Adoption of Internet technology at one establishment may complement or substitute adoption at another. As noted above, if establishments are substitutes and adoption costs vary by location, then urban density will be stronger for establishments in multi-establishment firms. We proceed in two stages. First, we present estimates in which we control for multi-establishment status, and allow the effects of location to vary for multi-

[30] These data are unweighted.

establishment firms. Second, we explicitly introduce into the model the adoption decisions of other establishments within the same firm.

First, we show the estimates where belonging to multi-establishment organization is a characteristic of an establishment. Multi-establishment status does predict adoption. The results in Tables 7a and 7b support the hypothesis that the effects of greater population size and density are larger (more positive) if the establishment is part of a multi-establishment firm. Equivalently, the effects of smaller population size and density are greater for establishments in single-establishment firms. This is particularly true for enhancement technologies.

For participation, interaction of a multi-establishment dummy with size and density causes many of the population variables that are not interacted to have a stronger economic and statistical significance. In other words, non-interacted population variables have a more negative impact than they had before. This suggests that the presence of multi-establishment firms may have slightly softened the effects of the global village theory in Table 4.

Column 1 of Table 7b shows that when a multi-establishment dummy is interacted with MSA dummies the non-interacted medium and large MSA dummies now both have a statistically significant marginal effect of –1.3% and –1.0% respectively. There is no statistically or economically significant effect in the interaction terms themselves. When multi-establishment is interacted with population, there is a statistically significant but small marginal impact of 0.042% per 100,000 person increase in population.[31] The effect of population itself has a more negative impact than in Table 4.[32] Similar results hold using population density rather than size.

[31] This calculation is not shown in the table.
[32] These results do not reflect any collinearity between mutli-establishment and urban areas. Multi-establishment dummies had a statistically and economically significant impact in the baseline regressions in

The economic significance of multi-establishment is larger for enhancement. As expected, interactions of multi-establishment dummies with population size and density weaken the positive effects of non-interacted population variables. In many cases coefficients become smaller or less statistically significant. However, in all the columns the interaction terms suggest that the effects of population size and density are greater for multi-establishment firms.

The economic effects can be quite substantial. Columns 5 and 8 suggest that establishments located in large MSAs or densely populated counties and part of multi-establishment firms are 2.2% to 2.3% more likely to adopt enhancement than stand-alone establishments in the same locations. These marginal effects are sizable compared to enhancement rates of 13%. The marginal effects of interacting multi-establishment with MSA population (in column 6) are statistically significant, but only 0.025% per 100,000 person increase in population.

Robustness checks: The discussion above ignores potential simultaneity bias arising from establishment-level analysis of adoption decisions in multi-establishment firms. We examined the robustness of our results in tables 4 and 7 by extending model (1) to include variables capturing the behavior of other establishments within the same firm. In particular, we added variables measuring the percentage and total of other establishments within the same firm adopting the dependent variable (e.g., participation or enhancement). Because these variables are likely to be correlated with unobserved factors affecting the decision to adopt participation and enhancement, we also used nonlinear instrumental variable techniques. For instruments, we used average population and density of other

---

Table 4; the marginal effect was between –2.6% and –2.8%. Moreover, the correlation between multi-establishment status and location in an MSA positive but small (0.0427).

establishments in the same firm. These should be correlated with adoption decisions at the firm's other establishments, but not at the establishment of interest.

Appendix Tables A.2 and A.3 show the marginal effects of probit regressions with and without IV that add other establishment adoption decisions to the models in Tables 4 and 7. Both tables show that these robustness checks make little difference to the relationship between population density and Internet adoption. Table A.2a shows that variables capturing the percentage of establishments with participation and enhancement come in positive and significant in weighted probit regressions, however their significance disappears once we instrument in Table A.2b. We interpret the probit regressions without IV as picking up unobserved heterogeneity, which is ultimately eliminated in the IV probit regressions. The new variables have little effect on the population marginal effects. However, statistical significance is retained in the IV version of this model.

Similarly, Appendix Table A.3 shows that the inclusion of other establishments' decisions has almost no impact on the estimates of our location variables. The new variables have the same pattern as in Table A.2: positive and significant in probit regressions without IV, insignificant in probit regressions with IV. In sum, we conclude that simultaneity of establishment decisions do not significantly bias our results.

7. Conclusions

Our conclusions significantly add to the broader literature that examines the relationship between advances in information technology and agglomeration of economic activity.

7.1 Is the Internet a Substitute or Complement to Agglomeration?

These findings inform the discussions of whether new IT is a complement or substitute with cities/agglomeration. We noted that this relationship is likely to change over the short and long run. We conducted this investigation at an early stage of the diffusion of a new information technology. It was late enough to see how establishments reacted to the availability of something new, but it was too early to observe wholesale relocation of these medium and large establishments in reaction to the availability of technology. Our results show that the question about substitutes and complements should be framed within the context of a particular technology and a particular stage of technology diffusion. Indeed, our results show that the relationship between general purpose IT and agglomeration may even change over the short run, as the benefits and costs of successive "waves" of technologies may differ across industries and locations.

We infer that geographically isolated establishments appeared to enjoy higher gross benefits because of initially high inter-organizational communication costs and few communications substitutes. Because the technology behind participation was well developed and required little co-invention, there was little variation in costs across different geographic areas. By 2000 participation had spread geographically throughout the country. The relationship between participation and agglomeration was, if anything, one of substitutes.

In contrast, enhancement technologies were complex and required substantial co-invention to be used successfully. We infer that enhancement and urban areas were complements over a time period in which Internet technology was diffusing. Establishments located in urban areas benefited from a combination of factors, including thicker labor markets, availability of complementary inputs, and knowledge spillovers. Thus, adoption costs were lower in urban areas.

Although location played an important role in the diffusion of participation and enhancement, we showed that over the short run its role was secondary to that of industry composition. Among major urban areas we observed a large number of locations with little difference in the extent to which specific areas contributed, on the margin, to the adoption of frontier technology. There appears to be a sense in which many urban areas are meaningful substitutes for each other in their contribution to the use of advanced technology.

We are observing the early stages of diffusion of a new technology. Sufficient time has not yet passed to observe relocation in response to a technology's availability. Industry composition produces heterogeneous response and provides clues about likely establishment mobility. That is, it tells us whether some locations contain the combination of density and industry to make them vulnerable to partial geographic dissolution of economic activity to other locations.

The establishments most likely to stay put in urban locations are those with intense demand for frontier Internet technology. For these establishments, the global village has less pull because the urban location acts as a complement to their demand for advanced Internet services. These establishments come from traditionally information intensive industries, such as company headquarters, utilities, finance and insurance, professional and scientific services, electronics manufacturing, and wholesale trade. Indeed, we forecast that there will be incentives for establishments from such industries in non-urban locations to relocate into urban areas.

In contrast, the establishments most likely to relocate from urban locations are those establishments who participate in the Internet, but who do not have intense demand for complex Internet applications. For these establishments, the global village has meaning. Urban and non-urban locations are less differentiated. These establishments come from traditionally less information intensive industries, and include general merchandisers, contractors, waste managers, textile mills, and equipment supply dealers.[33] We forecast that these are the establishments most at risk in urban locations to reconsider their location to less dense areas.

Moreover, we do not want to preclude the possibility of movement between cities by establishments. Ultimately, locations within different large, medium, and small urban areas and rural areas are better substitutes than locations across different size classes. Our results support the notion that innovations in IT lessen the transaction (and, in the case of information goods, transportation) costs of locating far from customers and suppliers. However, establishments may still choose to locate within urban areas to take advantage of thicker labor markets or knowledge spillovers.

7.2 Urban/rural divides in Internet use?

In one prominent view the Internet lowered communication costs and broke down geographic boundaries between firms.[34] The Internet networked much pre-existing IT, by increasing the marginal value of IT capital and bringing about a large increase in the rate of economic advance. An opposite view argues that the Internet exacerbated geographic inequalities, diffusing disproportionately to urban areas with complementary technical and knowledge resources.[35] Economic advance might have occurred, but it was concentrated in only a few locations. Our research shows that the latter arguments do not stand up to empirical analysis.

[33] For these and related findings about industries with a high propensity to adopt advanced applications, see Forman, Goldfarb and Greenstein (2002).
[34] The technology correspondent for the *Economist*, Cairncross (1997, p. 1), argues that "The death of distance as a determinant of the cost of communicating will probably be the single most important force shaping society in the first half of the next century."
[35] Articulation of this view is particular prominent in the National Telecommnications Information Administration's "Falling Through The Net" series (NTIA 1995, 1998, 1999, 2000).

Any references to a digital divide must be heavily qualified. The whole effect in any location is the sum of distinctly different marginal contributions. Once we control for variations in the benefits to Internet adoption across industries, there is some support for the global village theory in simple applications, such as participation, suggesting that the Internet may indeed help break down geographic and communication barriers. In enhancement, there remains some support for an urban density theory of Internet adoption even once we control for industry composition.

Nevertheless, within the more complex nexus of technologies defined by enhancement, assertions about a digital divide find only weak support. The pattern for enhancement is quite understandable as an economic matter. First, in smaller MSAs and rural areas, skew could arise from the thin technical labor markets alone. This would drive up costs of operating facilities employing Internet technology. Because the investment is linked to competitive settings, multi-establishment organizations, if they had a choice, would implement new business processes in more hospitable settings in major urban areas. Second, this reasoning also suggests that preexisting multi-establishment organizations would hesitate to open their own complex Internet facilities in rural areas until the costs are lower. Either case would lead to more use of enhancement in major urban areas.

Discussions about a "digital divide" between urban and rural areas are problematic because of systematic differences in industry and labor market composition across large urban and rural areas. More worrying from a public policy perspective is if Internet use remains concentrated in a few areas. However, this concern also does not stand up to empirical analysis. Participation and enhancement use remains widespread in large urban areas, while low levels of use in some smaller areas were driven by an industry composition highly concentrated in laggard industries. For this reason alone, public policy designed to equate Internet usage across regions may be inappropriate and would face high hurdles.

We are particularly skeptical of worries about the concentration of economic benefits from the diffusion of the Internet. There are over a dozen different industries comprised of thousands of establishments—from printing to information to finance to warehousing—intensively using the Internet. By the end of 2000 simple Internet technology had become a common facet for almost all United States business operations. Such a common factor can determine countrywide growth, but cannot determine the outcome of regional competitive advantage within the country, nor serve to exacerbate inequalities. That is, when many areas are comparatively similar in terms of use of the GPT and co-inventive activities, they compete on similar terms. All regions will improve at comparatively the same rate. Regional rivalry then will be determined by the less common factors (Furman, Porter, and Stern 2002), such as use of the Internet for enhancing business computing.

We conclude that research focused on concentration or digital divides—heretofore a central concern of the literature on Internet geography—is a misleading basis for framing the analysis of the use of Internet technology in business. Policies for regional development in most places should devote attention to the factors that are rare and possibly complementary to the use of the Internet for competitive advantage (e.g., such as immobile skilled labor, see Feldman 2002, Kolko, 2002). The concerns about low growth are real for the very few areas in which adoption lags due to location-specific factors; however, in the majority of cases adoption lags due to the composition of industry. In the majority of locations, therefore, policy aimed at increasing adoption will be misguided.

References

Atrostic, Barbara, K., and John Gates. 2001. U.S. Productivity and Electronic Business Processes in Manufacturing. Working Paper #01-11. Center for Economic Studies, US Census Bureau.

Atrostic, Barbara K., John Gates, and Ron Jarmin. 2000. Measuring the Electronic Economy: Current Status and Next Steps.. Working Paper #00-10. Center for Economic Studies, US Census Bureau.

Atrostic, Barbara, K., and Sang V. Nguyen. 2002. Computer Networks and U.S. Manufacturing Plant Productivity: New Evidence from the CNUS Data. Working Paper #02-01. Center for Economic Studies, US Census Bureau.

Bresnahan, Timothy, and Shane Greenstein. 1996. Technical Progress in Computing and in the Uses of Computers. *Brookings Papers on Economic Activity, Microeconomics,* 1-78.

Bresnahan, Timothy, and Shane Greenstein. 2001. The economic contribution of information technology: Towards comparative and user studies, *Journal of Evolutionary Economics,* v11, pp 95-118.

Bresnahan, Timothy, and Manuel Trajtenberg. 1995. General Purpose Technologies: 'Engines of Growth'? *Journal of Econometrics.* 65 (1): 83-108.

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt. 2002. Information Technology, Work Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics* 117 (February): 339-376.

Cairncross, Frances, 1997. *The Death of Distance.* Cambridge, MA: Harvard University Press.

Castells, Manuel, 2002. The Internet Galaxy, Reflections on the Internet, Business and Society, Oxford University Press.

Census. 2001. Home Computers and Internet Use in the United States: August 2000. Survey Release from the United States Department of Commerce. See www.census.gov/epcd/www/ebusiness.htm.

Census. 2002. Detailing Tabulations of Manufacturing E-Business Process Use in 2000. Survey Release from United States Department of Commerce. See www.census.gov/epcd/www/ebusiness.htm.

Charles, S., M. Ives, and A. Leduc. 2002. Embracing E-Business: Does Size Matter? Research Report, Statistics Canada.

Cherry, Barbara, Alan H. Hammond, and Steve Wildman. 1999. *Making Universal Service Policy: Enhancing the Process Through Multidisciplinary Evaluation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Compaine, Benjamin. 2001. *The Digital Divide: Facing a Crisis or Creating a Myth?* Cambridge, MA: MIT Press.

Downes, T.; Greenstein, S., 2002, “Universal Access and Local Internet Markets in the U.S.” *Research Policy*, 31, pp. 1035-1052.

Feldman, Maryann P., 2002. The Internet Revolution and the Geography of Innovation. Forthcoming in *International Social Sciences Review Journal*.

Forman, Chris. 2002. The Corporate Digital Divide: Determinants of Internet Adoption. Mimeo. Carnegie Mellon University.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2002. Digital Dispersion: An Industrial and Geographic Census of Commercial Internet Use. NBER Working Paper #9287.

Furman, Jeffrey L., Michael Porter, and Scott Stern. 2002. The Determinants of National Innovative Capacity. *Research Policy* 31(6): 899-933.

Glaeser, E., and Gaspar, J., 1997, “Information Technology and the Future of Cities.” *Journal of Urban Economics*, 43, pp. 136-156.

Greenstein, S. M., 2001, “Commercialization of the Internet: The Interaction of Public Policy and Private Actions,” in: A. Jaffe, J. Lerner and S. Stern (Editors), *Innovation, Policy and the Economy* (MIT Press, Cambridge, MA).

Hitt, Lorin and Erik Brynjolfsson. 1997. Information Technology and Internal Firm Organization: An Exploratory Analysis *Journal of Management Information Systems* 14 (Summer): 1-101.

Holmes, Thomas. 2002. The Role of Cities: Evidence from the Placement of Sales Offices. Federal Reserve Bank of Minneapolis Staff Report #298, January.

Hubbard, Tom. 2000. “The Demand for Monitoring Technologies: The Case of Trucking.” *Quarterly Journal of Economics* 116 (May): 533-560.

Ireland, N. and Stoneman, P. 1986. Technological diffusion, expectations and welfare. Oxford Economic Papers 38, 283-304.

Kitchen, R., and Dodge, Martin. 2001. 'Placing' cyberspace: why geography still matters. *Information Technology, Education and Society*, 1(2), 25-46.

Kolko, Jed, 2002, “Silicon Mountains, Silicon Molehills, Geographic Concentration and Convergence of Internet Industries in the US,” *Information Economics and Policy* 14(2), 211-232.

Kraemer, Kenneth, Jason Dedrick, and Debora Dunkle. 2002. E-Commerce: A Mile Wide and an Inch Deep. Working Paper, Center for Research on Information Technology and Organizations (CRITO), University of California, Irvine.

Krugman, Paul, 1991, *Geography and Trade*. Cambridge: MIT Press.

Malone, Thomas, Joanne Yates, and Robert Benjamin. 1987. Electronic Markets and Electronic Hierarchies. *Communications of the ACM* 30 (June): 484-497.

Mesenbourg, Thomas. 2001. Measuring Electronic Business: Definitions, Underlying Concepts, and Measurement Plans, mimeo released on http://www.census.gov/epcd/www/ebusines.htm.

Moss, Mitchell L., and Anthony M. Townsend, 1997. Tracking the net: using domain names to measure the growth of the Internet in US cities. *Journal of Urban Technology* 4(3): 47-60.

National Telecommunications Information Administration, US Department of Commerce, 2001. Falling Through the net: Toward Digital Inclusion.

---, 2000. "Advanced Telecommunications in Rural America, The Challenge of Bringing Broadband Service to All Americans." See http://www.ntia.doc.gov/reports.html.

---, 1998. "Falling Through the Net II: New Data on the Digital Divide." http://www.ntia.doc.gov/reports.html.

---, 1997, "Falling Through the Net: Defining the Digital Divide." http://www.ntia.doc.gov/reports.html.

---, 1995. "Falling Through the Net: A Survey of the 'Have Nots' in Rural and Urban America." http://www.ntia.doc.gov/reports.html.

Porter, Michael, 2000, "Strategy and the Internet," *Harvard Business Review.*

Premkumar G., 2000. "Rural Business and Telecommunications Technologies," in (eds.) Korsching, Peter F., Patricia C. Hipple and Eric A. Abbott. 2000. *Having All the Right Connections: Telecommunications and Rural Viability.* New York: Praeger.

Premkumar, G. and Margaret Roberts, 1999, "Adoption of New Information Technologies in Rural Small Business." *Omega - International Journal of Management Science*, 27, 467-484.

Rogers, Everett, 1995. *The Diffusion of Innovations.* Fourth Edition. New York: Free Press.

Saxenian, AnnaLee, 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Cambridge, MA: Harvard University Press.

Sinai, Todd and Joel Waldfogel, 2000, "Geography and the Internet: Is the Internet a Substitute or a Complement for Cities?" Working Paper, Wharton School of Business, University of Pennsylvania, Philadelphia.

Strover, Sharon, Michael Oden, and Nobuya Inagaki (2002), "Telecommunications and Rural Economics: Findings from the Appalachian Region," in (eds.) Lorrie Faith Cranor and Shane Greenstein, *Communications Policy and Information Technology: Promises, Problems, Prospects.* MIT Press.

Varian, Hal, Robert E. Litan, Andrew Elder, and Jay Shutter, 2001. The Net Impact Study. www.netimpactstudy.com.

Whinston, Andrew, Anitesh Barua, Jay Shutter, and Brant Wilson. Measuring the Internet Economy; available from http://www.internetindicators.com/jan_2001.pdf.

Zooks, Matthew, 2000a, "Internet Metrics: Using Hosts and Domain Counts to Map the Internet Globally," *Telecommunications Policy,* 24, 6/7.

Zooks, Matthew, 2000b, "The Web of Production: the Economic Geography of Commercial Internet Content Production in the United States," *Environment and Planning*, 32: 411-26.

Table 1
National Internet Adoption Rates (in percentages)

	Weighted Average	Unweighted Average
Participation	88.6%	80.7%
Enhancement	12.6%	11.2%
Enhancement & Experimenting with Enhancement	23.2%	18.1%

Table 2a: Participation Among Metropolitan Statistical Areas with Over One Million People

Avg. Rank	City	Avg. Rate	Std. Err. Rate	Marg. Rank	Marg. Coef.	Std. Err. Coef.	Obs.	Population
1	San Francisco--Oakland--San Jose, CA	96.4%	0.4%	7	0 (base)	N/A	2135	7,039,362
2	Denver--Boulder--Greeley, CO	95.9%	0.7%	43	-0.067	0.027	940	2,581,506
3	Cleveland--Akron, OH	94.8%	0.6%	23	-0.038	0.021	1099	2,945,831
4	Seattle--Tacoma--Bremerton, WA	93.9%	0.5%	3	0.025	0.015	1012	3,554,760
5	Salt Lake City--Ogden, UT	93.5%	0.8%	6	0.007	0.019	535	1,333,914
6	San Antonio, TX	93.3%	0.8%	1	0.035	0.021	395	1,592,383
7	Providence--Fall River--Warwick, RI--MA	93.0%	1.2%	24	-0.038	0.032	290	1,188,613
8	Grand Rapids--Muskegon--Holland, MI	93.0%	0.7%	4	0.012	0.021	503	1,088,514
9	Minneapolis--St. Paul, MN--WI	92.7%	0.5%	10	-0.011	0.017	1411	2,968,806
10	Los Angeles--Riverside--Orange County, CA	92.5%	0.4%	38	-0.061	0.017	4099	16,373,645
11	Kansas City, MO--KS	92.2%	0.6%	21	-0.035	0.025	753	1,776,062
12	Austin--San Marcos, TX	92.1%	0.7%	2	0.033	0.026	344	1,249,763
13	Dallas--Fort Worth, TX	92.1%	0.5%	36	-0.058	0.019	1720	5,221,801
14	Portland--Salem, OR--WA	92.1%	0.6%	5	0.009	0.019	776	2,265,223
15	Houston--Galveston--Brazoria, TX	91.7%	0.6%	17	-0.032	0.018	1413	4,669,571
16	Phoenix--Mesa, AZ	91.6%	0.7%	13	-0.022	0.018	988	3,251,876
17	Raleigh--Durham--Chapel Hill, NC	91.6%	0.9%	9	-0.004	0.028	398	1,187,941
18	Columbus, OH	91.5%	0.9%	28	-0.048	0.025	574	1,540,157
19	Milwaukee--Racine, WI	91.5%	0.7%	14	-0.023	0.023	855	1,689,572
20	San Diego, CA	91.5%	0.7%	32	-0.053	0.023	738	2,813,833
21	Detroit—Ann Arbor--Flint, MI	91.4%	0.6%	42	-0.067	0.021	1621	5,456,428
22	Indianapolis, IN	91.3%	0.8%	22	-0.036	0.024	646	1,607,486
23	Greensboro--Winston-Salem--High Point, NC	91.1%	0.9%	18	-0.032	0.024	570	1,251,509
24	Atlanta, GA	90.9%	0.6%	40	-0.064	0.024	1426	4,112,198
25	Miami--Fort Lauderdale, FL	90.9%	0.7%	35	-0.057	0.020	1010	3,876,380
26	Charlotte--Gastonia--Rock Hill, NC—SC	90.7%	0.9%	46	-0.083	0.029	618	1,499,293
27	Boston—Worcester--Lawrence, MA--NH--ME--CT	90.6%	0.5%	12	-0.022	0.015	2231	5,819,100
28	Chicago--Gary--Kenosha, IL--IN—WI	90.5%	0.4%	27	-0.047	0.016	3431	9,157,540
29	New York—Northern NJ--Long Island, NY--NJ--CT--PA	90.5%	0.4%	30	-0.050	0.015	4775	21,199,865
30	Washington--Baltimore, DC--MD—VA--WV	90.4%	0.5%	20	-0.034	0.017	2222	7,608,070
31	Philadelphia-Wilmington-Atlantic City, PA—NJ--DE--MD	90.3%	0.5%	16	-0.031	0.017	1745	6,188,463
32	Rochester, NY	90.3%	1.0%	19	-0.033	0.028	373	1,098,201
33	Hartford, CT	90.2%	0.9%	15	-0.024	0.027	500	1,183,110
34	Oklahoma City, OK	90.2%	1.1%	8	-0.002	0.024	339	1,083,346
35	Memphis, TN--AR--MS	90.0%	1.0%	26	-0.045	0.027	437	1,135,614
36	Louisville, KY--IN	89.9%	1.0%	25	-0.044	0.027	448	1,025,598
37	Cincinnati--Hamilton, OH--KY—IN	89.7%	0.8%	41	-0.066	0.024	772	1,979,202
38	St. Louis, MO--IL	89.7%	0.7%	11	-0.020	0.020	936	2,603,607
39	Pittsburgh, PA	89.1%	0.8%	34	-0.056	0.023	727	2,358,695
40	Buffalo--Niagara Falls, NY	88.5%	1.1%	31	-0.051	0.030	393	1,170,111
41	Tampa--St. Petersburg--Clearwater, FL	88.4%	0.9%	33	-0.054	0.021	812	2,395,997
42	Jacksonville, FL	87.6%	1.3%	47	-0.094	0.032	373	1,100,491
43	Las Vegas, NV--AZ	87.2%	1.2%	48	-0.106	0.030	417	1,563,282
44	Sacramento--Yolo, CA	87.0%	1.2%	45	-0.070	0.034	427	1,796,857
45	Norfolk--Virginia Beach--Newport News, VA--NC	86.9%	1.2%	49	-0.110	0.032	374	1,569,541
46	New Orleans, LA	86.0%	1.1%	37	-0.06	0.031	386	1,337,726
47	West Palm Beach--Boca Raton, FL	85.9%	1.2%	29	-0.049	0.029	299	1,131,184
48	Orlando, FL	85.5%	1.0%	44	-0.067	0.025	622	1,644,561
49	Nashville, TN	84.6%	1.1%	39	-0.062	0.028	466	1,231,311

Table 2b: Enhancement Among Metropolitan Statistical Areas with Over One Million People

Avg. Rank	City	Avg. Rate	Std. Err. Rate	Marg. Rank	Marg. Coef.	Std. Err. Coef.	Obs.	Population
1	Denver—Boulder--Greeley, CO	18.3%	1.3%	3	0.016	0.015	940	2,581,506
2	San Francisco--Oakland--San Jose, CA	17.0%	0.9%	15	0 (base)	N/A	2135	7,039,362
3	Salt Lake City--Ogden, UT	16.7%	1.7%	6	0.013	0.017	535	1,333,914
4	Minneapolis--St. Paul, MN--WI	15.9%	1.0%	10	0.003	0.012	1411	2,968,806
5	Houston--Galveston--Brazoria, TX	15.7%	1.0%	11	0.003	0.012	1413	4,669,571
6	Atlanta, GA	15.4%	1.0%	26	-0.008	0.011	1426	4,112,198
7	Oklahoma City, OK	15.4%	2.0%	2	0.020	0.021	339	1,083,346
8	Dallas--Fort Worth, TX	15.3%	0.9%	19	-0.003	0.011	1720	5,221,801
9	San Antonio, TX	15.3%	1.9%	4	0.013	0.020	395	1,592,383
10	Portland--Salem, OR--WA	15.1%	1.3%	5	0.013	0.019	776	2,265,223
11	Providence--Fall River—Warwick, RI--MA	14.9%	2.2%	7	0.010	0.024	290	1,188,613
12	Austin--San Marcos, TX	14.7%	1.9%	27	-0.009	0.016	344	1,249,763
13	Cleveland--Akron, OH	14.7%	1.2%	21	-0.004	0.014	1099	2,945,831
14	Tampa--St. Petersburg—Clearwater, FL	14.6%	1.3%	8	0.009	0.015	812	2,395,997
15	Memphis, TN--AR--MS	14.5%	1.8%	14	0.002	0.021	437	1,135,614
16	Seattle--Tacoma--Bremerton, WA	14.5%	1.2%	16	-0.002	0.012	1012	3,554,760
17	Hartford, CT	14.4%	1.6%	25	-0.008	0.016	500	1,183,110
18	San Diego, CA	14.3%	1.3%	23	-0.005	0.014	738	2,813,833
19	Cincinnati--Hamilton, OH—KY--IN	14.2%	1.3%	24	-0.005	0.014	772	1,979,202
20	Washington--Baltimore, DC—MD—VA--WV	14.2%	0.8%	22	-0.005	0.010	2222	7,608,070
21	Chicago--Gary--Kenosha, IL—IN--WI	14.1%	0.7%	17	-0.002	0.009	3431	9,157,540
22	Rochester, NY	14.1%	1.9%	18	-0.003	0.018	373	1,098,201
23	Boston--Worcester--Lawrence, MA--NH--ME--CT	13.9%	0.8%	20	-0.004	0.011	2231	5,819,100
24	Detroit--Ann Arbor--Flint, MI	13.8%	0.9%	39	-0.016	0.010	1621	5,456,428
25	Kansas City, MO--KS	13.7%	1.3%	35	-0.014	0.013	753	1,776,062
26	Raleigh--Durham--Chapel Hill, NC	13.7%	1.7%	31	-0.012	0.017	398	1,187,941
27	Pittsburgh, PA	13.6%	1.3%	13	0.003	0.015	727	2,358,695
28	Indianapolis, IN	13.6%	1.4%	41	-0.019	0.014	646	1,607,486
29	Charlotte—Gastonia--Rock Hill, NC--SC	13.6%	1.5%	29	-0.010	0.014	618	1,499,293
30	West Palm Beach--Boca Raton, FL	13.6%	2.0%	12	0.003	0.025	299	1,131,184
31	Los Angeles--Riverside—Orange County, CA	13.5%	0.6%	37	-0.015	0.008	4099	16,373,645
32	Miami—Fort Lauderdale, FL	13.5%	1.1%	33	-0.013	0.011	1010	3,876,380
33	New York--Northern NJ--Long Island, NY-NJ-CT-PA	13.5%	0.6%	36	-0.015	0.008	4775	21,199,865
34	Philadelphia-Wilmington-Atlantic City, PA--NJ--DE--MD	13.3%	0.9%	44	-0.021	0.009	1745	6,188,463
35	St. Louis, MO--IL	13.2%	1.2%	28	-0.009	0.013	936	2,603,607
36	Louisville, KY--IN	13.2%	1.6%	9	0.006	0.024	448	1,025,598
37	Columbus, OH	13.0%	1.5%	30	-0.011	0.018	574	1,540,157
38	Buffalo—Niagara Falls, NY	12.9%	1.7%	42	-0.019	0.014	393	1,170,111
39	Phoenix—Mesa, AZ	12.4%	1.1%	34	-0.014	0.012	988	3,251,876
40	Greensboro—Winston-Salem—High Point, NC	12.2%	1.4%	43	-0.020	0.015	570	1,251,509
41	Grand Rapids--Muskegon—Holland, MI	12.0%	1.5%	47	-0.031	0.012	503	1,088,514
42	New Orleans, LA	11.9%	1.7%	40	-0.018	0.016	386	1,337,726
43	Milwaukee--Racine, WI	11.7%	1.2%	38	-0.016	0.014	855	1,689,572
44	Nashville, TN	11.7%	1.5%	32	-0.012	0.016	466	1,231,311
45	Jacksonville, FL	11.3%	1.7%	48	-0.034	0.014	373	1,100,491
46	Sacramento--Yolo, CA	11.8%	1.6%	1	0.041	0.050	427	1,796,857
47	Norfolk--Virginia Beach—Newport News, VA--NC	10.8%	1.7%	45	-0.021	0.017	374	1,569,541
48	Orlando, FL	10.5%	1.3%	46	-0.027	0.012	622	1,644,561
49	Las Vegas, NV--AZ	9.0%	1.4%	49	-0.043	0.012	417	1,563,282

Table 3
Average Adoption by Size of Metropolitan Statistical Area

Population	Average Participation	Standard Error	Average Enhancement	Standard Error	Number of Areas
Non-MSA	85.1%	0.1%	10.6%	0.2%	49
<250,000	75.5%	0.2%	9.9%	0.3%	143
250,000-1 million	84.9%	0.2%	11.2%	0.3%	116
> 1 million	90.4%	0.1%	14.7%	0.2%	57

Table 4
Population Variables
(Standard errors in parentheses)

		Participation				Enhancement			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A.	Small MSA	0.0095				0.0198			
		(0.0285)				(0.0350)			
Coefficients from (Weighted) Probit Regressions	Medium MSA	-0.0491				0.0449			
		(0.0227)*				(0.0265)+			
	Large MSA	-0.0262				0.0632			
		(0.0201)				(0.0228)**			
	MSA Population		-3.95e-09	1.36e-08			-8.80e-10	2.10e-08	
			(3.80e-09)	(1.02e-08)			(3.24e-09)	(1.15e-08)+	
	MSA Population Squared			-1.94e-15				-2.39e-15	
				(1.13e-15)+				(1.24e-15)+	
	Medium-Low Density				-0.0170				0.0275
					(0.0194)				(0.0224)
	Medium-High Density				-0.0282				0.0860
					(0.0200)				(0.0244)**
	High Density				-0.0177				0.0577
					(0.0224)				(0.0224)*
	Log Likelihood	-33470.6	-33472.1	-33469.3	-33473.5	-28694.7	-28696.9	-28693.1	-28688.2
	Pseudo R^2	0.2252	0.2252	0.2252	0.2251	0.0593	0.0592	0.0593	0.0595

		Participation				Enhancement			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
B.	Small MSA	0.0021				0.0035			
		(0.0064)				(0.0062)			
Marginal Effects from (Weighted) Probit Regressions	Medium MSA	-0.011				0.008			
		(0.0052)*				(0.0048)+			
	Large MSA	-0.0058				0.0110			
		(0.0045)				(0.0039)**			
	MSA Population		-8.89e-10	3.06e-09			-1.54e-10	3.67e-09	
			(8.56e-10)	(2.30e-09)			(5.67e-10)	(2.02e-09)+	
	MSA Population Squared			-4.37e-16				-4.18e-16	
				(2.55e-16)+				(2.17e-16)+	
	Medium-Low Density				-0.00385				0.00485
					(0.00440)				(0.00399)
	Medium-High Density				-0.00639				0.0154
					(0.00456)				(0.00450)**
	High Density				-0.00400				0.0103
					(0.00508)				(0.00406)*

Notes:
standard errors are in parentheses.
(1) non-MSA is the base for these regressions
(2) & (3) Since no meaningful population data was available for non-MSA areas, we include a "rural area" dummy variable in each of these regressions. The population and density variables were interacted with (1-RURAL). Therefore the coefficients on the population variables do not include non-MSA areas.
(4) low density is the base for these regressions. One quarter of the observations fit into each density type.
+significant at 90% confidence level
*significant at 95% confidence level
**significant at 99% confidence level

Table 5
Average and Median Location Effects, by Type of Location

Type	N	Median Participation Marginal Effect	Average Participation Marginal Effect	Std Dev Participation Marginal Effect	Median Enhancement Marginal Effect	Average Enhancement Marginal Effect	Std Dev Enhancement Marginal Effect
Rural	49	-0.029	-0.0292	0.0486	-0.020	-0.0135	0.0274
Small MSA	130*	-0.0225	-0.0271	0.0772	-0.018	-0.00708	0.0495
Medium MSA	95	-0.046	-0.0535	0.0579	-0.012	-0.0111	0.0313
Large MSA	48	-0.0445	-0.0397	0.0324	-0.008	-0.00652	0.0150

*N=127 for enhancement because 3 small MSAs perfectly predicted non-adoption.

Table 6
Contribution of Industry and Location to Explaining Adoption Decisions

	Participation		Enhancement	
	Pseudo R^2	Log Likelihood	Pseudo R^2	Log Likelihood
Full model	0.2339	-33093.4	0.0672	-28443.4
No MSA Dummies	0.2251	-33475.0	0.0591	-28701.4
No NAICS dummies	0.1526	-36604.2	0.0347	-29434.6

Note: Cities defined by CMSA.

Table 7a
Population Variable Coefficients from (weighted) Probit Regressions,
Includes Multi-Establishment/Population Interactions
(Standard errors in parenthesis)

	Participation				Enhancement			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Multi-est. dummy	-.1511	-0.1661	-0.1661	-0.1509	-0.0578	-0.0014	-0.0038	-0.0463
	(0.0339)**	(0.0211)**	(0.0210)**	(0.0289)**	(0.0404)	(0.0210)	(0.0212)	(0.0334)
Small MSA	0.0247				0.0013			
	(0.0430)				(0.0492)			
Medium MSA	-0.0571				0.0317			
	(0.0321)+				(0.0342)			
Large MSA	-0.0463				0.0134			
	(0.0279)+				(0.0284)			
Small MSA ×	-0.0267				0.0532			
Multi-est. dummy	(0.0570)				(0.0696)			
Medium MSA ×	0.0196				0.0436			
Multi-est. dummy	(0.0446)				(0.0538)			
Large MSA ×	0.0442				0.1234			
Multi-est. dummy	(0.0378)				(0.0447)**			
MSA Population		-1.27e-08*	-3.07e-09			-4.96e-09	2.07e-08	
		(5.16e-09)	(9.37e-09)			(4.08e-09)	(1.03e-08)*	
MSA Population			-1.09e-15				-2.92e-15	
Squared			(1.02e-15)				(1.13e-15)**	
MSA Population ×		1.86e-08**	1.82e-08**			1.45e-08	1.44e-08	
Multi-est. dummy		(6.38e-09)	(6.29e-09)			(5.91e-09)*	(6.01e-09)*	
Medium-Low				-0.0222				-0.0024
Density				(0.0280)				(0.0294)
Medium-High				-0.0212				0.0574
Density				(0.0286)				(0.0341)+
High Density				-0.0619				0.0059
				(0.0323)+				(0.0289)
Medium-low				0.0132				0.0760
density × Multi-est.				(0.0382)				(0.0450)+
Medium-high				-0.0112				0.0738
density × Multi-est.				(0.0386)				(0.0465)
High density ×				0.0961				0.1238
Multi-est.				(0.0417)*				(0.0435)**
Log Likelihood	-33467.9	-33463.5	-33462.5	-33465.8	-28686.8	-28695.9	-28689.2	-28681.9
Pseudo R^2	0.2253	0.2254	0.2254	0.2253	0.0596	0.0593	0.0595	0.0597

Notes:
All regressions include dummy variables for 3-digits NAICS, month that data was collected, and whether it was a multi-establishment firm. Employment and Employment squared were also included as controls. Population was measured at the MSA level.
(1) non-MSA is the base for these regressions
(2) & (3) Since no meaningful population data was available for non-MSA areas, we include a "rural area" dummy variable in each of these regressions. The population and density variables were interacted with (1-RURAL). Therefore the coefficients on the population variables do not include non-MSA areas.
(4) Low density is base for these regressions. One quarter of the observations fit into each density type.
+significant at 90% confidence level
*significant at 95% confidence level
**significant at 99% confidence level

Table 7b
Population Variable Marginal Effects from (weighted) probit regressions,
Includes multi-establishment effects
(Standard errors in parenthesis)

	Participation				Enhancement			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Multi-est. dummy	-0.0343	-0.0377	-0.0377	-0.0342	-0.0101	-0.0002	-0.0007	-0.0081
	(0.0077)**	(0.0048) **	(0.0048) *	(0.0066)	(0.0070)	(0.0037)	(0.0037)	(0.0058)
Small MSA	0.0055				0.0002			
	(0.0095)				(0.0086)			
Medium MSA	-0.0131				0.0056			
	(0.0075)+				(0.0061)			
Large MSA	-0.0104				0.0023			
	(0.0062)+				(0.0050)			
Small MSA ×	-0.0061				0.0096			
Multi-est. dummy	(0.0132)				(0.0130)			
Medium MSA ×	0.0043				0.0078			
Multi-est. dummy	(0.0099)				(0.0098)			
Large MSA ×	0.0099				0.0224**			
Multi-est. dummy	(0.0083)				(0.0084)			
MSA Population		-2.85e-09	-6.91e-10			-8.68e-10	3.63e-09	
		(1.17e-09)*	(2.11e-09)			(7.15e-10)	(1.81e-09)*	
MSA Population			-2.44e-16				-5.12e-16	
Squared			(2.30e-16)				(1.98e-16)**	
MSA Population ×		4.19e-09	4.09e-09			2.53e-09	2.52e-09	
Multi-est. dummy		(1.44e-09)**	(1.42e-09) **			(1.03e-09)*	(1.05e-09)*	
Medium-Low				-0.0050				-0.0004
Density				(0.0064)				(0.0051)
Medium-High				-0.0048				0.0102
Density				(0.0065)				(0.0062)+
High Density				-0.0141				0.0010
				(0.0075)+				(0.0051)
Medium-low				0.0030				0.0138
density × Multi-est.				(0.0085)				(0.0085)
Medium-high				-0.0025				0.0134
density × Multi-est.				(0.0088)				(0.0087)
High density ×				0.0208				0.0229
Multi-est. dummy				(0.0087)*				(0.0085)**

Notes:
All regressions include dummy variables for 3-digits NAICS, month that data was collected, and whether it was a multi-establishment firm. Employment and Employment squared were also included as controls. Population was measured at the MSA level.
(1) non-MSA is the base for these regressions
(2) & (3) Since no meaningful population data was available for non-MSA areas, we include a "rural area" dummy variable in each of these regressions. The population and density variables were interacted with (1-RURAL). Therefore the coefficients on the population variables do not include non-MSA areas.
(4) Low density is base for these regressions. One quarter of the observations fit into each density type.
+significant at 90% confidence level
*significant at 95% confidence level
**significant at 99% confidence level

Figure 1
Comparison by City Size of Marginal Effects for Participation

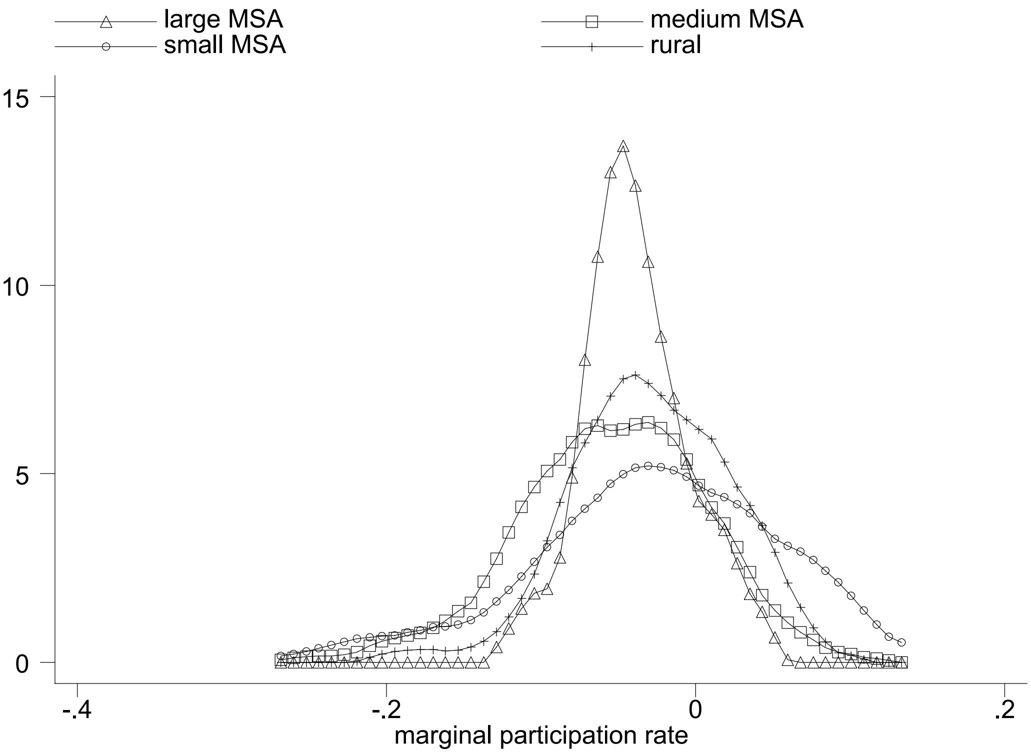

Figure 2
Comparison by City Size of Marginal Effects for Enhancement

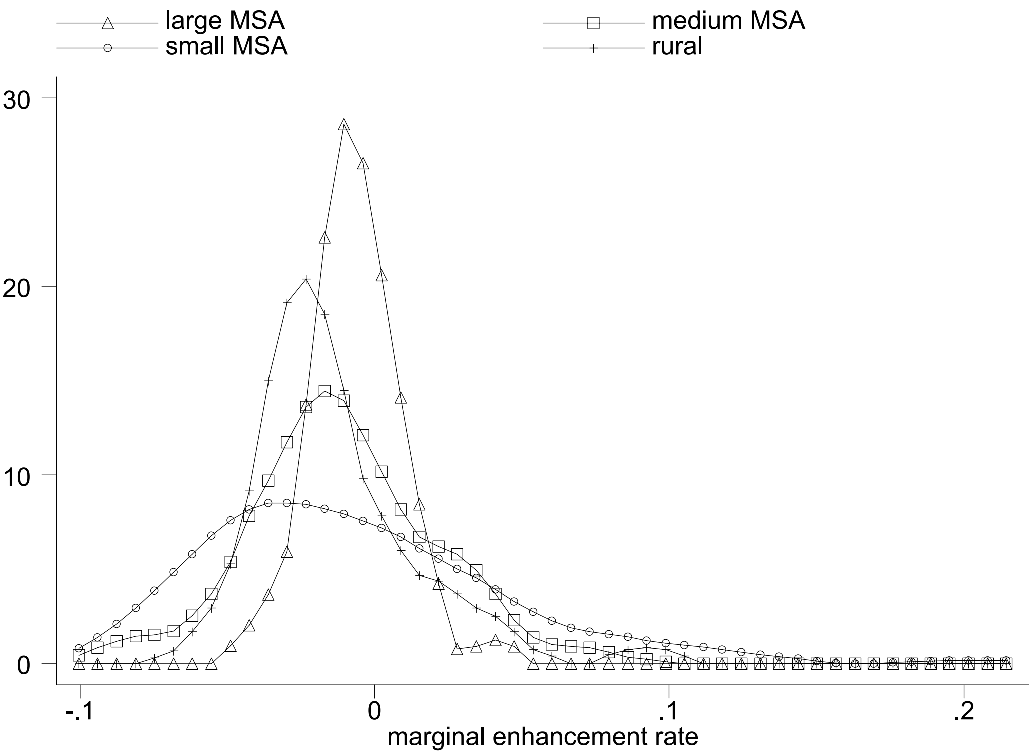

Figure 3
Differences in Industry Marginal Effects for Participation

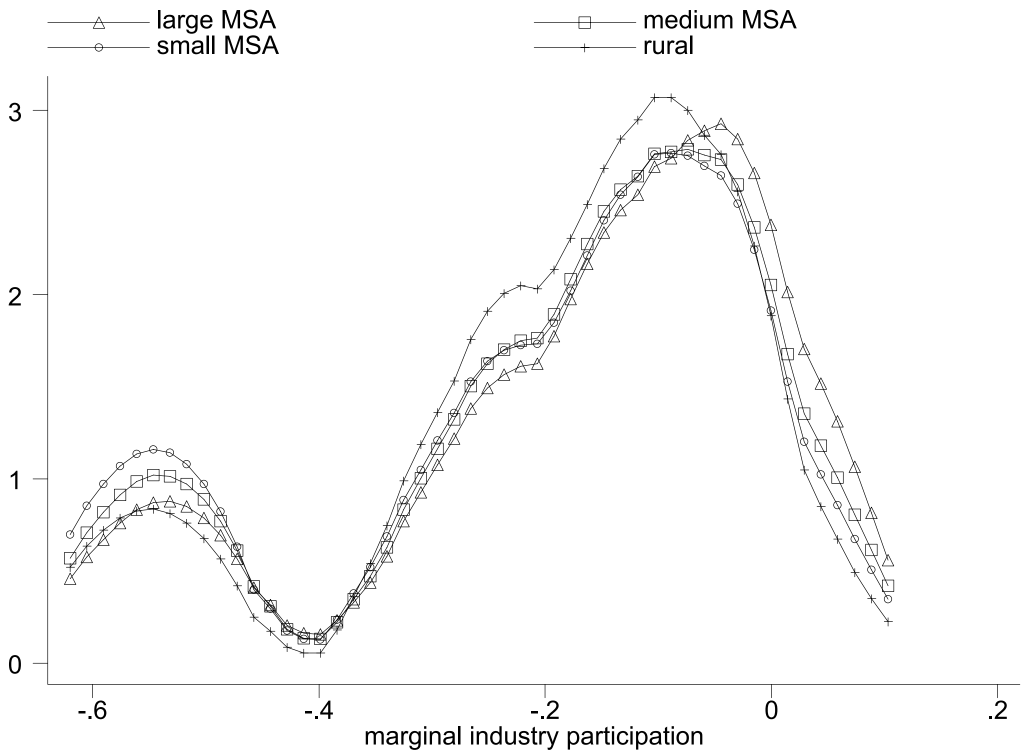

Figure 4
Differences in Industry Marginal Effects for Enhancement

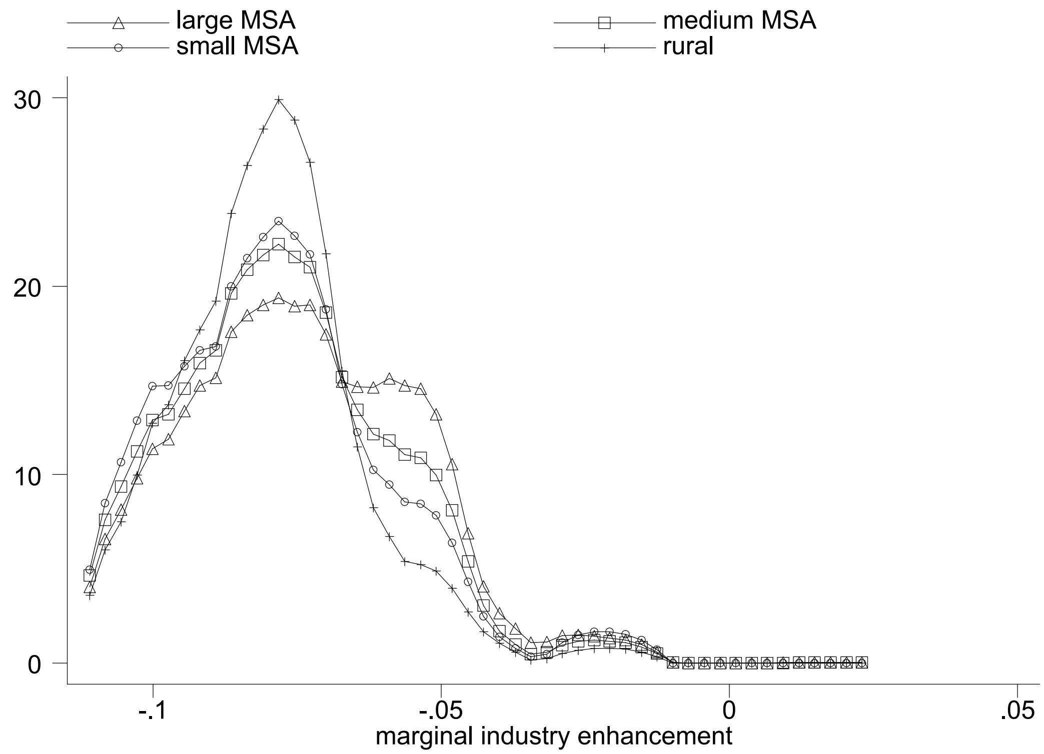

Table A1: Harte Hanks Sample Versus the Census of Business Establishments

	Sample	Census
# establishments with over 100 employees	86,879	168,372
% MSA	82.5%	86.7%
% CMSA	37.2%	42.5%
% >500 employees given have 100 employees	12.7%	10.6%
% Northeast	17.7%	19.6%
% Midwest	27.9%	25.5%
% South	34.8%	34.0%
% West	19.6%	21.0%
% Agriculture, Forestry, Fishing and Hunting (NAICS=11)	0.2%	0.1%
% Mining (NAICS=21)	0.6%	0.5%
% Utilities (NAICS =22)	0.8%	0.8%
% Construction (NAICS =23)	2.9%	4.1%
% manufacturing (NAICS =31,32,33)	27.9%	20.8%
% Wholesale Trade (NAICS =42)	6.0%	4.8%
% Retail Trade (NAICS =44,45)	17.1%	14.7%
% Transportation & Warehousing (NAICS =48, 49)	2.9%	3.1%
% Media, Telecommunications and Data Processing (NAICS =51)	3.7%	3.7%
% Finance and Insurance (NAICS =52)	4.5%	4.6%
% Real Estate and Rental and Leasing (NAICS =53)	0.5%	1.0%
% Professional, Scientific and Technical Services (NAICS =54)	5.2%	5.0%
% Management of Companies and Enterprises (NAICS =55)	0.3%	3.2%
% Administrative and Support and Waste Management and Remediation Services (NAICS =56)	2.7%	10.2%
% Educational Services (NAICS =61)	0.01%	1.2%
% Health Care and Social Assistance (NAICS =62)	16.7%	12.8%
% Arts, Entertainment, and Recreation (NAICS =71)	1.6%	1.5%
% Accommodation and Food Services (NAICS =72)	5.5%	5.1%
% Other Services (except Public Administration) (NAICS =81)	0.9%	2.2%

Table A.2

Population Variable Marginal Effects from Probit Regressions in Table 4, Includes Percent Participation and Enhancement Adopters within Firm

	Model	Variable	Old Result	New Result
A. Weighted probits without IV	Add percentage other establishments adopting participation to column (1)	Small MSA	0.0021	0.0021
		Medium MSA	-0.0110*	-0.0112*
		Large MSA	-0.0058	-0.0063
		Pct shallow	N/A	0.2401**
	Add percentage other establishments adopting enhancement to column (5)	Small MSA	0.0035	0.0038
		Medium MSA	0.0080+	0.0081+
		Large MSA	0.0110**	0.0108**
		Pct deep	N/A	0.1026**
	Add percentage other establishments adopting participation to column (4)	Medium-Low Density	-0.0039	-0.0039
		Medium-High Density	-0.0064	-0.0069
		High Density	-0.0040	-0.0052
		Pct shallow	N/A	0.1637**
	Add percentage other establishments adopting enhancement to column (8)	Medium-Low Density	0.0049	0.0049
		Medium-High Density	0.0154**	0.0152**
		High Density	0.0103*	0.0099*
		Pct deep	N/A	0.1024**
B. Unweighted probits with IV*	Add percentage other establishments adopting participation to column (1) (instrument using average population)	Small MSA	0.0032	0.0025
		Medium MSA	-0.0072+	-0.0075+
		Large MSA	-0.0045	-0.0051
		Pct shallow	N/A	0.0193
	Add percentage other establishments adopting enhancement to column (5) (instrument using average population)	Small MSA	0.0095*	0.0097*
		Medium MSA	0.0077*	0.0078*
		Large MSA	0.0129**	0.0128**
		Pct shallow	N/A	0.0336
	Add percentage other establishments adopting participation to column (4) (instrument using average density)	Medium-Low Density	-0.0019	-0.0013
		Medium-High Density	-0.0012	-0.0010
		High Density	-0.0027	-0.0028
		Pct shallow	N/A	0.1538**
	Add percentage other establishments adopting enhancement to column (8) (instrument using average density)	Medium-Low Density	0.0044	0.0042
		Medium-High Density	0.0167**	0.0171**
		High Density	0.0110**	0.0114**
		Pct shallow	N/A	-0.1604

Notes:
Table compares results of probit regressions with and without variables measuring behavior of other establishments within the same firm.
"Old" coefficients are different because probits are unweighted.
+significant at 90% confidence level
*significant at 95% confidence level
**significant at 99% confidence level

Table A.3
Population Variable Marginal Effects from Probit Regressions in Table 7, Includes Percent Participation and Enhancement Adopters within Firm

	Model	Variable	Old Result	New Result
A. Weighted probits without IV	Add percentage other establishments adopting participation to column (1)	Small MSA	0.0055	0.0042
		Medium MSA	-0.0131+	-0.0138+
		Large MSA	-0.0104+	-0.0111+
		Small MSA * Multi-est dummy	-0.0061	-0.0099
		Medium MSA * Multi-est dummy	0.0043	0.0054
		Large MSA * Multi-est dummy	0.0099	0.0098
		Multi-est dummy	-0.0343**	-0.1510**
		Pct shallow	N/A	0.1636**
	Add percentage other establishments adopting enhancement to column (5)	Small MSA	0.0002	0.0005
		Medium MSA	0.0056	0.0059
		Large MSA	0.0023	0.0029
		Small MSA * Multi-est dummy	0.0096	0.0096
		Medium MSA * Multi-est dummy	0.0078	0.0071
		Large MSA * Multi-est dummy	0.0224**	0.0205*
		Multi-est dummy	-0.0101	-0.0219**
		Pct deep	N/A	0.1015**
	Add percentage other establishments adopting participation to column (4)	Medium-Low Density	-0.0050	-0.0057
		Medium-High Density	-0.0048	-0.0058
		High Density	-0.0141+	-0.0145*
		Medium-Low Density * Multi-est dummy	0.0030	0.0042
		Medium-High Density * Multi-est dummy	-0.0025	-0.0016
		High Density * Multi-est dummy	0.0208*	0.0195*
		Multi-est dummy	-0.0342	-0.1505**
		Pct shallow	N/A	0.1634**
	Add percentage other establishments adopting enhancement to column (8)	Medium-Low Density	-0.0004	-0.0001
		Medium-High Density	0.0102+	0.0107+
		High Density	0.0010	0.0016
		Medium-Low Density * Multi-est dummy	0.0138	0.0131
		Medium-High Density * Multi-est dummy	0.0134	0.0117
		High Density * Multi-est dummy	0.0229**	0.0207*
		Multi-est dummy	-0.0081	-0.0199**
		Pct deep	N/A	0.1014**

	Model	Variable	Old Result	New Result
B. Unweighted probits with IV*	Add percentage other establishments adopting participation to column (1) (instrument using average population)	Small MSA	0.0050	0.0041
		Medium MSA	-0.0076	-0.0081
		Large MSA	-0.0072	-0.0077+
		Small MSA * Multi-est dummy	-0.0032	-0.0032
		Medium MSA * Multi-est dummy	0.0011	0.0011
		Large MSA * Multi-est dummy	0.0056	0.0049
		Multi-est dummy	-0.0315***	-0.0311+
		Pct shallow	N/A	0.0150
	Add percentage other establishments adopting enhancement to column (5) (instrument using average population)	Small MSA	0.0065	0.0064
		Medium MSA	0.0029	0.0026
		Large MSA	0.0017	0.0013
		Small MSA * Multi-est dummy	0.0090	0.0091
		Medium MSA * Multi-est dummy	0.0137+	0.0140+
		Large MSA * Multi-est dummy	0.0288**	0.0299**
		Multi-est dummy	-0.0125*	-0.0037
		Pct shallow	N/A	-0.0904
	Add percentage other establishments adopting participation to column (4) (instrument using average density)	Medium-Low Density	-0.0016	-0.0021
		Medium-High Density	0.0023	0.0013
		High Density	-0.0113*	-0.0115*
		Medium-Low Density * Multi-est dummy	-0.0004	0.0018
		Medium-High Density * Multi-est dummy	-0.0065	-0.0042
		High Density * Multi-est dummy	0.0175**	0.0177**
		Multi-est dummy	-0.0315**	-0.1316**
		Pct shallow	N/A	0.1384**
	Add percentage other establishments adopting enhancement to column (8) (instrument using average density)	Medium-Low Density	-0.0017	-0.0024
		Medium-High Density	0.0052	0.0041
		High Density	0.0000	-0.0015
		Medium-Low Density * Multi-est dummy	0.0159*	0.0173**
		Medium-High Density * Multi-est dummy	0.0279**	0.0326**
		High Density * Multi-est dummy	0.0279**	0.0336**
		Multi-est dummy	-0.0113*	0.0156
		Pct shallow	N/A	-0.2599

Notes:
Table compares results of probit regressions with and without variables measuring behavior of other establishments within the same firm.
"Old" coefficients are different because probits are unweighted.
+significant at 90% confidence level
*significant at 95% confidence level
**significant at 99% confidence level